Biochemical and structural analysis of the <italic>Streptomyces coelicolor Dps</italic> proteins.

Hitchings, Matthew D

Biochemical and Structural Analysis of the *Streptomyces coelicolor Dps* proteins

By
MATTHEW D. HITCHINGS

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

Swansea University
Prifysgol Abertawe

Institute of Life Science, College of Medicine
Swansea University

September 2013

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

The three DNA protection proteins from starved cells (Dps) of *Streptomyces coelicolor* are members of the mini-ferritin super family. Considered to be of major importance to stress response systems in microorganisms, Dps proteins can aid microbial survival in extreme conditions. The *S. coelicolor* Dps proteins are not only induced in response to stress in a stimulus-dependent manner, but dual regulation allows these proteins to play a role in bacterial cell division; influencing condensation of nucleoids during spore formation.

This study investigates the structural and functional properties of the *Sc*Dps proteins and finds multiple ways in which the homologs differ. The proteins present differential assembly and ferroxidase characteristics. DpsA and DpsC are both capable of self-assembly to form the representative Dps dodecamers *in vivo* and *in vitro*. DpsB yields an assembly that is limited to a dimeric state, but retains ferroxidase activity. Subsequently all three proteins afford oxidative protection to DNA without evidence of a direct Dps/DNA interaction. The crystal structures of DpsA and DpsC highlight the divergence between the two proteins. A peculiarity of the phylogenetically distinct DpsC protein is its enhanced dodecameric stability, which is attributed to its "extremophile-like" structural architectures. These are comprised of enhanced interface surface areas, which are combined with an increase in abundance of salt bridges and hydrogen bonds that represent a strategy often adopted by extremophiles, but in this case are present in a mesophilic organism.

Together with mutational studies, structural analysis points toward the importance of the varying length N and C terminal tails for assembly and stability of the dodecamers. The various tail less mutants display dysfunctional abilities to correctly oligomerise into dodecamers. The disturbed assembly of the DpsA tail less variants has been utilised as a means of modulating and controlling multimeric assembly. Together with the creation of functionalised DpsA subunits, the ability to form hetero-oligomeric assemblies capable of forming an iron oxide core has led to these proteins being used as a platform to develop functionalised paramagnetic nanoparticles.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed:

Date: 27/5/2014

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed:

Date: 27/5/2014

STATEMENT 2

I hereby give consent for my thesis, if successful, may be made available for inter-library loan or photocopying (subject to the law of copyright), and that the title and summary may be made available to outside organisations only after expiry of a bar.

Signed:

Date: 27/5/2014

Contents

Acknowledgements

My sincere gratitude goes to all who have helped, advised, supported and provided encouragement over the past few years. Many thanks to Dr Ricardo Del Sol for all his time and effort in providing his help, guidance and imparting his extensive research knowledge. I wish to thank Prof Paul Dyson for his advice and support whilst working within his research group.

I am grateful to Mr Meirwyn Evans for his work in keeping the shelves stocked. And a special mention goes to Dr Paul Facey for his support and valuable help. I would also like to thank all my lab colleagues, present and previous, for making lab life enjoyable.

Finally, I wish to express my thankfulness to my family for all their encouragement.

Diolch!

Tables

Figures

Abbreviations

1D	One-dimensional
2D	Two-dimensional
APBS	Adaptive Poisson-Boltzmann Solver
APS	Ammonium persulphate
Å	Angstrom
Bfr	Bacterioferritin
BN	Blue-native

bp	Base pair(s)
BSA	Bovine Serum Albumin
CBD	Cell binding domain
C. diff	*Clostridium difficile*
CDS	Coding domain sequence
CSS	Complex formation Significance Score
DCT	Deleted C-tail
DNA	Deoxyribose nucleic acid
DNT	Deleted N-tail
dNTP	Deoxynucleoside Triphosphate
Dps	DNA protection proteins from starved cells/ DNA binding proteins from starved cells
ds	double stranded
DTM	Deteted Tail Mutant
EDTA	Ethylene Diamine Tetraacetic acid
Eq.	Equation
ESSP	Electrostatic surface potential
FOC	Ferroxidase Centre
Ftn	Ferritin
g	Gram(s)
h	Hour(s)
Hb(s)	Hydrogen bond(s)
HRP	Horse radish peroxidise
IMAC	Immobilised metal ion affinity chromatography
IPTG	Isopropyl-β-*D*-thiogalactoside
kb	kilobase
kDa	Kilodalton(s)
kcal/mol	kilocalorie per mole
L or l	Litre

mCh	mCherry
MCS	Multi cloning site
mg	milligram
min	Minute(s)
µ	Micro
mM	millimolar
MW	Molecular Weight
ng	nanogram
nm	Nanometre
PAGE	Polyarcylamide gel electrophoresis
PDB	Protein Data Bank
pI	Isoelectric point
PISA	Proteins, Interfaces, Structures and Assemblies
PCR	Polymerase Chain Reaction
PVDF	polyvinylidene difluoride
RMSD	Root Mean Square Deviation
ROS	Reactive Oxygen Spcecies
rpm	Revolutions per minute
sec/s	Second(s)
Sb(s)	Salt bridge(s)
SDM	Site directed mutagenesis
SDS	Sodium Dodecyl Sulphate
TBE	Tris Borate EDTA
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Tris	Tris(hydroxymethyl)amino-methane
v/v	volume/volume
V	Volt(s)
W	Watt(s)

Chapter 1

Introduction

1.1 Introduction

Stochastic fluctuations of environmental conditions in which microorganisms such as bacteria live have been, in part, the driving force of natural selection that has led to the evolution of the great diversity of organisms within the bacterial Kingdom (Bennett and Lenski, 1997). Through millions of years, bacteria have evolved sophisticated and often elaborate mechanisms to efficiently detect and combat potentially damaging environmental changes and stresses such as nutrient limitations, temperature instabilities, pH changes and osmotic pressures to name but a few (Groves and Lucana, 2010). Although the external environment can provide a harsh assault on a bacterium's struggle for survival, it is by no means the only source of stress. The continuous generation of reactive oxygen species (ROS) as by-products of routine metabolic processes such as aerobic respiration can also potentially lead to oxidative damage (Mailloux and Harper, 2011). It has been documented that as a result of reducing molecular oxygen, the respiratory chain enzymes can produce up to 87% of hydrogen peroxide created within the *E. coli* cell (González-Flecha and Demple, 1995). The potency of hydrogen peroxide as a damaging compound itself is limited, however it's potential to give rise to highly toxic and reactive hydroxyl radicals ($^{\bullet}OH$) and hydroxyl anions (OH^-) through Fenton chemistry (Eq. (1)) render it dangerous. The reaction involving the oxidation of transition metals such as ferrous iron by H_2O_2 (Luo et al., 1994) increases its harmful capabilities creating products capable of producing DNA strand breaks, peroxidation of lipids and damage to other intracellular macro molecules (Halliwell and Gutteridge, 1984).

$$Fe\,(II) + H_2O_2 \rightarrow Fe\,(III) + \bullet OH + OH^- \qquad (1)$$

As an indispensable element for almost all life forms, iron can be found playing numerous roles in many areas of biology, from the well-known constituent of the oxygen transporter haemoglobin (Uchida et al., 1970), to the use as a metal-based cofactor for the activity of ribonucleotide reductases aiding the formation of deoxynucleotides in both prokaryotes and eukaryotes (Cotruvo and Stubbe, 2012). Yet the very properties of iron that are lent to these diverse and ultimately essential functions, such as its range of oxidation states (-2 to +6) and the redox potential of the Fe (II)/Fe (III) couple, can also potentially lead to undesirable side effects. In the presence of oxygen, iron is commonly found in an oxidised state as toxic, insoluble ferric hydroxide precipitates (Briat, 1992) consequently limiting the availability of cellular free iron. In order to prevent iron precipitation, reducing agents within the cell are capable of converting uncomplexed ferric iron to its divalent ferrous form. However this initiates a cycle that favours generation of reactive oxygen species (ROS) through the process of the Haber-Weiss reaction (Harrison and Arosio, 1996). The incomplete reduction of O_2 generates a superoxide anion $_{\bullet}O_2^-$, and hydrogen peroxide H_2O_2, both of which are involved

in the iron oxidation/reduction cycle as shown in equations below; (2) – Fenton reaction, (3) – iron reduction reaction and the sum of these equations given in equation (4) – Haber-Weiss reaction.

$$Fe\ (II) + H_2O_2 \rightarrow Fe\ (III) + {}^{\bullet}OH + OH^- \quad (2)$$

$$O_2^{\bullet-} + Fe\ (III) \rightarrow O_2 + Fe\ (II) \quad (3)$$

$$H_2O_2 + O_2^{\bullet-} \rightarrow OH^- + {}^{\bullet}OH + O_2 \quad (4)$$

The oxidising potency of iron is not just a concern when within the cell; the equilibrium of polymerised complex iron to free Fe (III) is tipped against the favoured free iron. This causes a problem not with abundance, but with bioavailability (Ilbert and Bonnefoy, 2013). Bacteria have accordingly developed various strategies to maintain iron homeostasis and as a consequence microbial iron cycling plays a major role in various (bio) geochemical cycles that take place in most microbial niches from soils to marine environments (Kappler and Straub, 2005).

1.2 The Ferritin Superfamily

Ferritins are members of an extensive super-family of iron related proteins that are fundamental to the metabolism and control of cellular iron. The evolutionary conservation and ubiquitous nature to all three domains of life emphasises the importance of classical ferritin (Ftn) to nearly all organisms from *Homo sapiens* (Lawson et al., 1991) through to *Pyrococcus furiosus* (Matias et al., 2005). It also highlights the significant role that iron fulfils in biology (Andrews, 2010). This protein family can be divided into three sub-families; classical ferritins (Ftn), the haem binding Bactrioferritins (Bfr), and DNA protection proteins from starved cells (Dps) also known as DNA binding proteins from starved cells. Whilst there may be a divergence in their exact biological roles, architectures and phylogenetic distribution, they all share a common function; iron homeostasis. Not only are these proteins capable of acting as a bioavailable iron reserve; supplying the cell with the necessary concentration of iron as and when required. But in sequestering and detoxifying the free ferrous iron within a cell, these proteins also provide oxidative protection, safeguarding the cell and its content against oxidative damage (Abdul-Tehrani *et al.*, 1999).

1.2.1 Classical Ferritin

The classical ferritins (Ftn) are the archetypal members of the ferritin super-family and belong to the maxi-ferritin subgroup. Indeed they are the principal form of iron storage in most organisms. Forming tetracosameric protein conglomerates (24-mers) (Crichton and Declercq, 2010), the protein's subunits self-assemble to form a spherical protein cage with a hollow internal cavity. It is this cavity that allows the protein the ability of depositing and storing a polynuclear hydrous iron oxide core (Levi et al., 1989). In this state the iron is water soluble, non-toxic and remains bioavailable. There are a number of structural elements that provide the protein with the ability to carry out this process.

1.2.2 Subunit Structure – A Conserved Four Helix Bundle

A character conserved amongst all of the members of the super family is the structural motif of a four-helix bundle monomer (Figure 1.1). The Ftn monomer has a core four-helix bundle consisting of four long helices (A, B, C and D helices). In addition a series of loops, tails and an additional short fifth helix (E helix), located on a C-terminal extension, also compose the monomer. This small helix adjoins the end of the cylindrical-like bundle pointing towards the core at approximately 60° (Ha et al., 1999). Furthermore its importance to the control of assembly and ensuing stability of ferritin has been demonstrated by destabilisation of the structures upon its removal (Fan et al., 2009). The four main helices are maintained by amino acid side chain interactions; a number of hydrophobic and perhaps more importantly hydrophilic interactions. The preservation of the extensive numbers of intra-subunit hydrogen bonds is thought to significantly contribute to the hyperthermophilic properties of certain ferritins (Tatur et al., 2007). The non-helical regions of the proteins; the tail extensions at the terminal regions and the inter-helical loops are also important for the maintenance of the cylindrical like four helix bundle and of course form important features once the protein assembles.

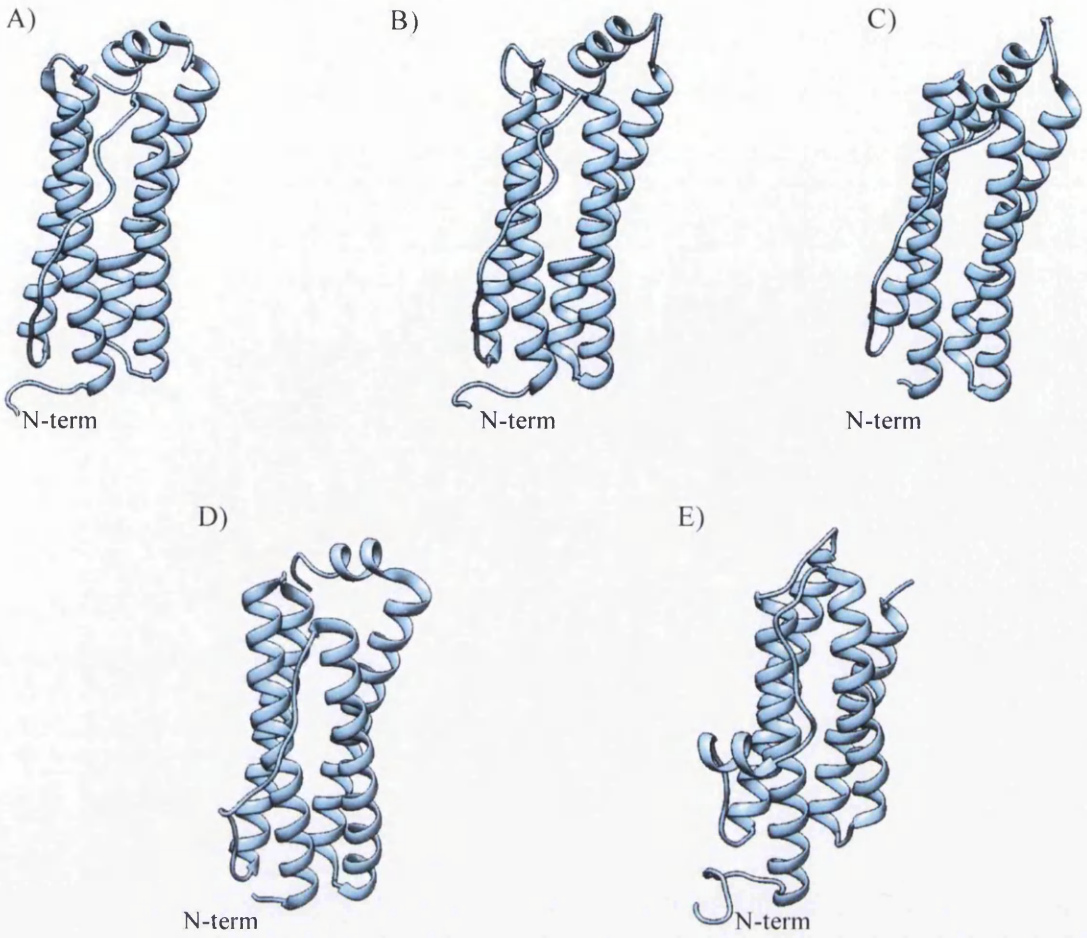


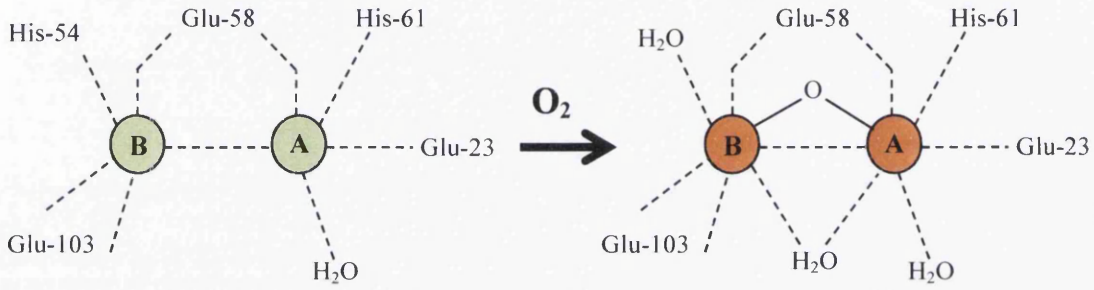
Figure 1.1: Monomeric ribbon diagrams of proteins belonging to the ferritin superfamily. (A) Human H chain ferritin (PDB ID 1FHA), (B) Human L chain ferritin (PDB ID 2FG8), (C) *C. jejuni* Ftn (PDB ID 1KRQ), (D) *E. coli* haem containing Bfr (PDB ID 2VXI) and (E) *E. coli* Dps (PDB ID 1DPS). Images created in UCSF Chimera.

In vertebrates, ferritin is typically found as isoferritin; a structure composed of two different subunits. The subunits, H (heavy) and L (light), have similar sequences (~55%) but despite their structural similarity, they remain distinct. Whilst the superposition generated root-mean-square-deviation (RMSD) between alpha carbon atoms of residues in H ferritin and L ferritin yields very low values, highlighting this structural similarity between proteins with only loop structures displaying slight divergences (Crichton and Declercq, 2010), they differ in their molecular weight; H chain ~ 21 kDa and L chain ~19 kDa (Munro et al., 1988). An interesting feature of mammalian ferritin is the tissue specific composition of the multimeric assemblies. The proportion of H to L subunits found within the 24-mers changes in a tissue dependent manner; human muscle ferritin has a ratio of 20-H to 4-L, while in liver tissue 3-H to 21-L is common (Harrison and Arosio, 1996). The reason behind varied chimeric assembly is the

functional differences found between the H and L chains. The core of the four-helix bundle affords a region where the largest difference between H and L subunits is found. It is this location that is associated with catalytic activity of ferritin. The H chain is endowed with an intra-subunit carboxylate-bridged di iron ferroxidase centre (FOC) (Lawson et al., 1991). The site conserved between Ftn, bacterial Ftn and Bfr is located within the four helix bundle and is responsible for the conversion of ferrous iron to ferric iron in an apolar environment (Wardeska et al., 1986). Each of the helices lend residues to coordinate iron often amounting to six residues in total and often display the metal binding motif *HxxE* (Rajasekaran et al., 2010). The donation of these residues is conserved not only between proteins, but also their helical location. The use of glutamates (Glu) and histidine residues (His) is preserved at the FOC of Ftn and Bfr and is thought to be a relic to early rubrerythrin ancestors that utilise these residues as iron binding ligands (Andrews, 2010). Glutamates are located on each of the four helices while the histidine residues that constitute the FOC are common only to the B and D helices (Andrews, 2010). Through these residues, the proteins preserve the ability to bind two iron atoms per FOC (Wahlgren et al., 2012).

The two iron binding sites, known as the A site and B sites have an almost symmetrical architecture (Schematic 1). The binding of Fe (II) iron is thought to be short lived; through its coupling with oxygen, peroxo-diferric intermediates have been detected after just 25 ms (Hwang et al., 2000). Because of this very little structural data has been generated in which the initial coordination environments are captured. Most structures soaked with ferrous iron display Fe (III) coordinated at site A and B bridged with oxygen. Many studies have found that site A is the preferred iron-binding site. Its mixed oxygen and nitrogen ligands renders it the first site to bind metal and thus in many structures the B site is left unoccupied (Bou-Abdallah et al., 2002). On the other hand mutations to residues at the A and B sites confirm that both are required for efficient and rapid oxidation. Mutations to A sites decrease the binding affinity for Fe (II), while modifications to the B site do not prevent the binding of Fe (II) to the A site but inhibit oxidation (Treffry et al., 1997, Bou-Abdallah et al., 2002). Furthermore, a recent study utilising X-ray data from flash frozen iron and redox stable metal soaked crystals have provided results which help to confirm a mechanism for the iron binding and oxidation process and emphasises the importance of the B site to iron oxidation turnover. Binding of metals at the FOC of frog ferritin is accompanied by the formation of a carboxylate bridge (Glu-58) between the two atoms. When structures containing the redox stable Cu (II) were compared to iron soaked crystals, a shortening in distance between atoms at the A and B sites from ~ 4.3 Å in Cu (II) crystals to ~ 3.1 Å in iron crystals was observed (Bertini et al., 2012). This suggests that oxidation occurs in parallel to a shortening of metal-to-metal distances. Furthermore large RMSD of the carboxylate bridge forming Glu-58 residue and a B site ligand His-54 point to the

simultaneous reorientation of the carboxylate bridge and detachment of iron from His-54 resulting in the rotation of Glu-58 and His-54 side chains. The source of the force postulated to close the gap between irons is the formation of a μ-oxo-bridged di-iron (III) complex (Bou-Abdallah, 2010). It is also hypothesised that the reoriented residues pave the way for the binding of new Fe (II) and a weakened protein/Fe (III) complex affinity provides a means for its translocation toward the cavity (Bertini et al., 2012). The oxidation process in frog ferritin is represented in the schematic 1, below.



Schematic 1: Structural model depicting binding of iron (II) at the diiron ferroxidase centre of frog H chain ferritin. Based on the structure of bound redox stable copper (II) (left), formation of an oxo-bridge (right) and reorientation of the Glu-58 Carboxylate Bridge is thought to contribute to the detachment of iron from His-54 upon oxidation to ferric iron. Diagram adapted from (Bertini et al., 2012).

While the H chain Ftn subunits catalyse the oxidation of ferrous iron, the L chain Ftn is involved with a complementary function involving the phases of nucleation and iron (III) core formation. As such the L chain does not possess an active FOC (Wang et al., 2006) and thus has no ferroxidase activity. Instead, the L chain four-helix bundle creates negatively charged diamond patches on the internal surface of the assembled 24-mer from a number of glutamate side chains. This region, located on the B helix at the dimer interface, serves as a ferrihydrite nucleation centre (Lawson et al., 1991) and is located to accelerate the translocation of iron from the FOC of the H chain to the cavity of the protein (Bou-Abdallah, 2010). It is this functional variation between subunits that promotes the differential hetero-assembly of the 24-mers in tissue dependant manners; H rich assemblies are associated with a metallo-scavenging roles due to their rapid oxidation capabilities, while L rich assemblies, which generally produce larger iron cores, are utilised as iron storage vessels (Chasteen and Harrison, 1999).

1.2.3 Multimeric Structure – A Spherical Protein Cage

Despite the variation in cage composition of vertebrate ferritins, the protein cage remains constant; measuring 12-13 nm in diameter with the internal cavity measuring around 7-8 nm and is capable of storing up to 4500 iron atoms (Dedman et al., 1992). The exact assembly pathway taken by the subunits to form the icositetramer is somewhat still an enigma; however proposed models exist based on experimental evidence. A notion that only stable dimers precede the formation of the assembly with a series of structural intermediates was explored with the use of intrinsic fluorescence, circular dichroism and protein cross-linking (Gerl and Jaenicke, 1987). However evidence pointed to the formation of a trimer fashioned from a dimer and monomer. Only then does the assembly proceed to form hexamers followed by dodecamers and finally the tetracosamer. Even if the exact pathway remains unclear, the consensus stands that dimer formation is almost certainly the foremost assembly (Stefanini et al., 1987), (Gerl and Jaenicke, 1987). Dimer formation occurs in an antiparallel manner; the E helices at the C-terminal of each subunit are positioned at alternative ends of the dimer producing symmetry with two-fold rotation. The interface is composed of the A and B helices of each subunit between which a series of hydrophobic and hydrophilic interactions occur rapidly ensuing the correct tertiary folding of each monomer. Due to these interactions, the dimer interface maintains a considerable amount of stability. Resistance of dimer dissociation has been observed in a range of treatments from low pH to incubation with denaturing agents such as urea (Linder et al., 1989). This evidence would advocate the assembly pattern with a series of intermediates of increasing dimers. Harsh physio-chemical environmental stimuli cannot break apart the two subunits that form a dimer yet re-establishment of the 24-mer that occurs upon restoration of physiological conditions indicates that formation of the 24-mer occurs from the amalgamation of 12 dimers.

Once assembled, the protein cage exhibits an octahedral symmetry (Granier et al., 2003) and contains a series of symmetry related subunit arrangements that create import functional features of ferritin (Figure 1.2). Apart from the twelve dimers with 2-fold symmetry, the 24-mer also contains two other symmetry related protein interfaces. Eight trimeric interfaces and six tetramers contribute to the assembly which ultimately dictates that each subunit forms six different inter-subunit interactions (Zhang and Orner, 2011).

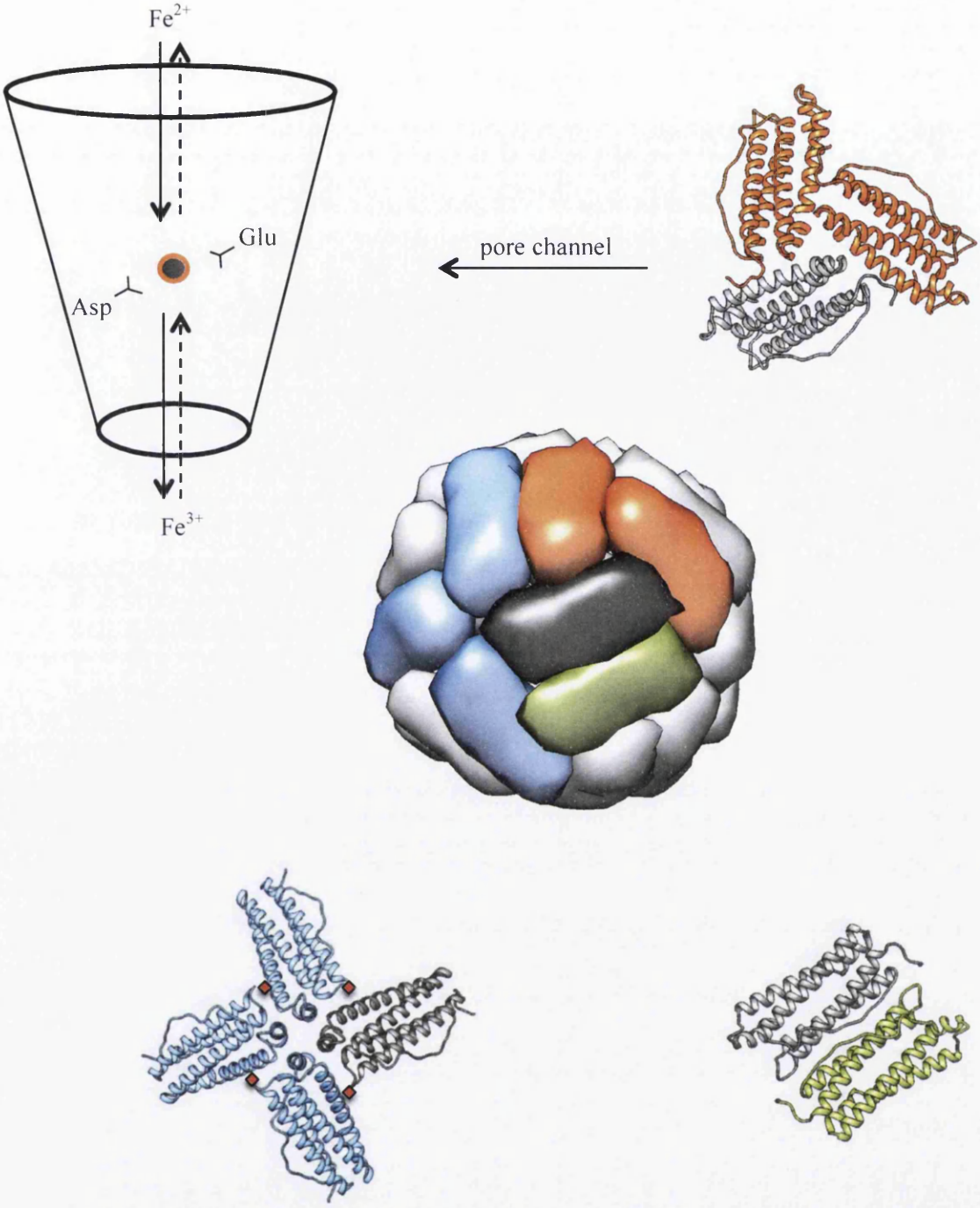


Figure 1.2: The biological unit multiscale model of bullfrog M ferritin (PDB ID 1MFR). The low-resolution surface for each monomer is accompanied by the ribbon structures displaying the three types of symmetry related interfaces created within the 24-mer assemblies. Blue – tetrameric interface (6 per 24-mer), orange – trimeric interface (8 per 24-mer) and green – dimer interface (12 per 24-mer). The dark grey subunit interacts with all six coloured monomers. Red squares represent approximate location of B-pores present in Bfr proteins. Images were created using UCFS Chimera 1.7rc. Also depicted is a schematic for iron transportation through the "funnel-like" three-fold pore utilising conserved acidic Asp and Glu residues for the coordination of iron.

The threefold interfaces creates funnel shaped channels which run through the protein shell, linking the external environment with the internal environment (Figure 1.2) (Ford et al., 1984). It is understood that this hydrophilic pore, approximately 12 Å in length, provides a route for molecules to move through the protein cage both into and out of the protein cavity. However the negative charge based method of translocation favours movement of positive molecules such as iron and excludes negatively charged diffusants (Yang and Chasteen, 1996). Highly conserved Asp and Glu residues line the internal region of the pore and are involved in coordinating iron, guiding its transportation toward the FOC. The importance of a negative electrostatic gradient for iron translocation and ferroxidase activity has been confirmed by mutagenesis studies to *E. coli* ferritin (*Ec*FtnA – PDB ID 1EUM). Point mutations to negative Asp-131 and Glu-134 channel residues abate ferroxidase activity by five- to seven-fold. However, a significantly slower rate of oxidation was observed for FOC mutations with preserved 3-fold pores. This indicates that the FOC is the site for iron oxidation (Treffry et al., 1993).

In contrast to the three-fold pores, the four-fold pores are lined with non-polar residues and are hydrophobic in nature. With narrower dimensions than the three fold pores, the four fold pores are predicted not to be involved in the coordination of divalent or monovalent ions (Takahashi and Kuyucak, 2003). Instead, the four-fold pore has been speculated to create a "proton wire" formed from hydrogen bonded water molecules within the pore's channel (Takahashi and Kuyucak, 2003). The "proton wire" could be used to transfer the protons formed during the oxidation process out of the cage and supplies a more efficient route than if transferred via the three-fold pore. However, with little conservation of residues within this region, the four-fold pore may have evolved protein/organism specific roles.

1.3 Bfr and Bacterial Ferritins

The two forms of maxi-ferritins found in prokaryotes share significant similarities with classical ferritins yet there are subtle but important differences. They both maintain the 24-mer assemblies composed of four helix bundle monomers, nevertheless despite structural similarity the sequences share very few conserved amino acids other than the important residues which constitute the FOC (Le Brun et al., 2010). Both bacterial Ftn and Bfr are considered analogous to the H-chain ferritin monomers of eukaryotic origin. However, as displayed by *Ec*FtnA, bacterial ferritins show a greater similarity in sequence to H chain ferritin than they do to Bfr proteins (Hudson et al., 1993). Oligomeric structural heterogeneity has been observed in bacterial Ftn and Bfr where genomes contain more than one gene encoding Ftn or Bfr (Moore et al., 1994) although the exact frequency at which this occurs within prokaryotes is unknown.

Crystal structures of a number of Ftn (Stillman et al., 2001), (Yao et al., 2011) and Bfr (Crow et al., 2009) (Gupta et al., 2009) confirm the conservation of the symmetry-related structural arrangements of subunits that create the pores into the cavity of the proteins; the three-fold pore and the four-fold pore. However the amino acid composition responsible for these regions is not preserved, producing divergence in function between Bfr and bacterial Ftn proteins. Classical Ftn proteins have functionally important negatively charged amino acids present at the three-fold pores (Treffry et al., 1993) and have crystal structures revealing the presence of divalent metals within these channels indicating their role in metal movement (Hempstead et al., 1997). Bfr structures display a variety of charges at this pore alternating between positive and negative in layers through the channel. The *M. smegmatis* Bfr (*Ms*-Bfr – PDB ID 3B3H) three-fold pore contains five layers of negative/positive charges lining the channel surface pointing to the ineffectiveness of this region to coordinate cations (Janowski et al., 2008). Furthermore the number of positive residues presented at the three-fold pore of the *E. coli* Bfr (*Ec*Bfr – PDB ID 1BCF) suggests that despite its hydrophilic nature, metal transportation is not likely to be a role for this region (Frolow et al., 1994).

In contrast to Ftn, Bfr four-fold channels do contain hydrophilic residues, albeit placed in layers with hydrophobic residues so that a pattern of hydrophobic external region is accompanied with a hydrophilic inner region. Even this patterning is not conserved within Bfr proteins. *Ms*-Bfr displays three layers consisting of a hydrophobic outer layer caused by side chains of leucine residues followed by a middle hydrophilic layer of glutamine. At the cavity end of the channel, arginine side chains form a positively charged inner layer (Janowski et al., 2008). The Bfr of *D. desulfuricans* (*Dd*Bfr – PDB ID 1NF6) only has two layers; an external nucleophilic threonine layer is coupled with a positive internal lysine layer (Macedo et al., 2003). Some evidence exists that suggest cations can be transported via asparagine and glutamates at the four-fold pore in the Bfr of *A. vinelandii* (*Av*-Bfr – PDB ID 1SOF) (Swartz et al., 2006), however many other Bfr present pores are devoid of cations in their crystal structures and as such the role of this pore remains unclear.

Another distinct characteristic possessed by Bfr and not by classical ferritin is the presence of pores other than the three- and four-fold pores known as the B site exhibited in Figure 1.2 (Macedo et al., 2003). This is another site where the functional conservation is indefinite and whilst its exact amino acid composition tends to differ between proteins, the B site is usually hydrophilic in nature and negatively charged. The structures of *Ms*-Bfr (Janowski et al., 2008) and *Av*-Bfr (Swartz et al., 2006) both display Mg^{2+} surrounded by Asp, Asn and Gln ligands at the B site. Given its closer locality to the FOC than the three- and four-fold pores, (~17 Å vs 21 Å and 32 Å respectively), it is considered plausible for this to serve as an iron pathway since its ability to coordinate divalent cations has been demonstrated.

The main difference that separates Ftn, both eukaryotic and prokaryotic from Bfr is the presence of non-covalently bound haem moieties associated with Bfr. The haem molecules are located at the interface between the monomers that create the symmetry related dimers and thus each 24-mer contains 12 haem moieties. Despite the difficulties encountered whilst preparing proteins with reproducible haem content, the haem groups are thought to be directly related to mineralisation and release of the iron core. The haem group is reported to accelerate core formation by facilitating electron transportation (Wong et al., 2012). Although it is not required for ferroxidase activity (Andrews et al., 1995), evidence depicting a reduced oxidation rate in the presence of the haem group has been reported (Wong et al., 2012). However, in contrast, an increase in rate and extent of iron release by a pathway in which the FOC is not involved has also been documented suggesting the haem to be important for iron reduction (Yasmin et al., 2011).

1.4 Dps proteins – Mini Ferritins with Diverse Functions

Dps are the third major group of the ferritin super family that constitute the mini-ferritins. First discovered in 1992, the most abundant *E. coli* stationary phase protein gained its name from the *in vivo* stable, non-specific interactions it forms with DNA in starved *E. coli* cells (Almirón et al., 1992). Furthermore the oxidative protection, which these proteins confer, was also described for the first time. It was not until the crystal structure of the *E. coli* Dps protein was resolved (*Ec*Dps – PDB ID 1DPS) that structural and functional homologies to ferritin were revealed (Grant et al., 1998). The monomeric folding pattern is analogous to ferritin, yielding a highly similar four-helix bundle. However while ferritin assembles into 24-mers, Dps monomers pack together to form smaller homo-dodecamers (12-mer) with 233-point symmetry (Grant et al., 1998). The external diameter of the Dps dodecamer is ~9 nm. The ability to store iron is also maintained by the formation of an internal cavity which although being smaller in dimensions compared to classical Ftn, at ~4-5 nm, it is capable of holding up to 500 iron atoms (Zhang and Orner, 2011). Iron can be oxidised and deposited within the cavity via the conserved Dps ferroxidase centres. However these sites are not located within the monomer, instead they are located at the interface between two antiparallel subunits that form the two-fold symmetry dimers; each monomer donating residues involved in iron co-ordination (Bellapadrona et al., 2010). A preference for the use of the more efficient oxidant hydrogen peroxide over oxygen as a ferrous iron oxidants means that in species that encode both Ftn and Dps, the Dps is primarily used for its oxidative protection role and not for iron storage (Chiancone and Ceci, 2010). These basic biochemical and biophysical attributes have attracted attention to the Dps family as being smaller alternatives to ferritin, and as such, provide

different properties to exploit for the production of novel biotechnological agents (Flenniken et al., 2009)

Since their initial discovery, the increasing number of sequenced microbial genomes has revealed an abundance of Dps homologues, almost ubiquitous within bacteria and common to many archaea. Subsequently, biochemical characterisation highlighting the structural and functional insights of the Dps continually grows and expands the knowledge base. Consequently the PDB database (http://www.rcsb.org) contains the crystal structures of Dps from 26 different bacterial species and 2 archaeal species. Within this array of organisms there is a range of ecological niches occupied; from the gastric mucosa infected by the human pathogen *H. pylori* (*HP*-NAP – PDB ID 1JI4) (Zanotti et al., 2002) (Lai et al., 2013) to high radiation environments colonised by *K. radiotolerans* (*Kr*Dps – PDB ID 4A25) (Ardini et al., 2013). Interestingly, despite the conservation of protein structure at tertiary and quaternary levels, the Dps family appear to have been adapted for cellular protection in a multitude of environments and accordingly exhibit a variety of diverse functions.

The multimeric assembly of the Dps proteins endows most of the members with the ferritin trait of iron sequestering. Oxidative protection of DNA and other macromolecules is mediated by the ability of many of the proteins to utilise hydrogen peroxide as an oxidant in addition to molecular oxygen during ferroxidation (Chiancone et al., 2004). Furthermore DNA binding ability is another recognised trait displayed among the group by many Dps. Despite the family being named after this property, not all Dps possess this ability. In addition to the standard "Ftn-like" features, Dps have also been implicated in a wide range of cellular stress responses; cold shock response (Nicodème et al., 2004), pH and metal stress (Tsou et al., 2008), host inflammatory response via neutrophil activation (Choli-Papadopoulou et al., 2011), biofilm formation (Theoret et al., 2012) and host colonisation/survival (Halsey et al., 2004), with further diverse roles continually being revealed. Of course it is the strong structure/function relationship that allows these proteins to perform their tasks, yet only slight deviations of just a few altered amino acids result in a unique functionality being generated.

1.4.1 Features of the Dps Four Helix Bundle Monomer

The Dps monomer is composed of a four-helix bundle (helices named; A, B, C and D) with an additional small helix (BC helix). The packing of these helices generally involves hydrophobic interactions (Franceschini et al., 2006). The four-helix bundle adopts an "up-down-down-up" arrangement allowing the fully folded helix pairs (A+B and C+D) to remain antiparallel. This also allows the N and C termini to be placed at opposing ends of the bundle (Gauss et al., 2006). The folding motif is accomplished by a long loop that stretches from the end of the second helix (B helix) to the beginning of the third helix (C helix), traversing the

bundle's entire length (Roy et al., 2004). Within the B to C loop is the location of the BC helix. This short helix orientated centrally and orthogonally to the four-helix bundle, is one of the main features that differentiate Dps from Ftn. Indeed the addition of a Bfr E helix to a Dps monomer produced assemblies with dimensions larger than that of the Dps dodecamer (Fan et al., 2009). In the same study, removal of the BC helix from the Dps did not disrupt assembly but reduced temperature stability. This is in contrast with the loss of the E helix from Bfr, which results in abolishment of oligomeric assembly above a dimeric state and suggests that the E helix is more influential in directing multimeric assembly of Bfr than the BC helix is in Dps. The BC helix has however been connected with the ability to strengthen a cubic crystalline lattice structure of the *HP*-NAP where each dodecamer interacts with 6 other dodecamers. The inter-dodecamer interactions (hydrogen bonds) occur between Lys-74 (BC helix) of one dodecamer and Ile-73 and Thr-76 (BC helix) of an adjacent dodecamer (Tsuruta et al., 2012). Figure 1.1 displays the similarity of the central four-helix bundles between mammalian Ftn, bacterial Ftn, Bfr and Dps. As expected, RMSD calculations reported in the literature confirm that Dps proteins yield the largest distance values between C^{α} atoms when aligned with Ftn proteins (Crichton and Declercq, 2010). It is not only the differences in the helical placement that cause a large RMSD, but varying loop and tail structures also contribute to the RMSD values.

1.4.2 Dps Multimeric Assembly – A Smaller Protein Cage

The multimeric assembly of Dps proteins produces structural arrangements that resemble those involved in creating the Ftn 24-mers. Instead of exhibiting the 234 symmetry found in ferritins, Dps create a two-fold symmetry axis and two sets of three-fold symmetry axis (23 point symmetry) allowing each subunit to interface with five other subunits. Along with the 6 dimer interfaces responsible for the establishment of the inter-subunit ferroxidase centre, there is also the formation of 4 "ferritin-like" trimer interfaces and 4 "Dps-like" trimer interfaces. The "ferritin-like" timer is so called because it not only resembles the arrangement of the three-fold symmetry related subunits of Ftn, but in forming a pore, it is also associated with iron translocation (Roy et al., 2004). The "Dps-like" trimer defines an interface common only to Dps; nevertheless it is reminiscent of the functional ambiguity of the Ftn four-fold symmetry axis. The "Dps-like" pore is diverse in amino acid composition and hence its exact function remains unclear throughout the family (Chiancone and Ceci, 2010). To understand how these interfaces contribute to Dps assembly, it is possible to consider each independently.

1.4.3 "Ferritin-like" Trimeric Interface – A Route for Iron Transport

The "ferritin-like" trimer is formed from the N-terminal regions of the subunits, creating a pore through the cage to the cavity of the protein. As previously discussed for Ftn,

this pore is involved in iron translocation. Thus due to the functional significance it is largely conserved among Dps. The channel opens on the surface of the dodecamer with diameters that range between approximately 9 Å and 17 Å based on Cα-Cα distances. The channel progresses in a funnel shape toward the cavity with a depth of approximately 10 Å. On the inner surface, the diameter decreases to ~ 4-11 Å (Haikarainen and Papageorgiou, 2010). There is however, the possibility of side chain rotational flexibility within these pore regions. This can provide a means to modulate the dimensions of the channel openings. This function can been seen in the Dps of *S. solfataricus* (SsDps – PDB ID 2CLB), where rotation of tyrosine side chains may facilitate the opening of the channel (Gauss et al., 2006). The Dps of *S. enterica* (SeDps – PDB ID 3AK8) creates a channel with an "hour glass" shape, where the outer opening of 14 Å is constricted to down to 1.7 Å before opening up at the cavity surface to 12 Å. Although it is thought that the three Asp-146 residues that create the constriction serve a functional purpose; replacing the water molecules which hydrate ferrous iron thus preparing the iron for oxidation by dehydration (Miyamoto et al., 2011a).

The N-terminal pore is by and large hydrophilic and is lined with acidic residues. This makes the region negatively charged under physiological conditions. Just like Ftn, it is thought that these residues create a negative electrostatic gradient; guiding iron through the protein shell and toward the FOC (Takahashi and Kuyucak, 2003). The *L. innocua* Dps (*Li*Dps – PDB ID 1QGH) has a pore lined with Asp-121, Asp-126 and Asp-130. The inner channel Asp-130 residue in particular modulates iron uptake and as a result influences the size of the mineralised core (Bellapadrona et al., 2009). This inner Asp residue (Asp-130 in *Li*Dps) is conserved in many Dps, and as documented in the table of pore sizes collated by Haikarainen and Papageorgiou (2010), all but one Dps presented an Asp residue at the cavity region of channel. *H. salinarum* was the exception in the list, placing a Glu-154 residue instead. In this case the pore is said to be closed by a series of residues (Glu-141, His-150, Arg-153, and Glu-154) and would require major conformational changes to open. However the three Glu-154 present at the cavity side of the three-fold pore alternatively create an iron nucleation centre (Zeth et al., 2004). Iron translocation in *Halobacterium salinarium* (*Hs*Dps – PDB ID 1MOJ) is thus suggested to take place via an unconventional set of 12 pores composed of three translocation sites (T1, T2 and T3). The pores are located close to the dimer interface and are thought to be an adaptation in favour of its extreme environment.

The Dps2 from *D. radiodurans* (*Dr*Dps2 – PDB ID 2C2J) displays another unusual N-terminal pore, orthologous to the *Hs*Dps. A positively charged Lys residue is located on the external surface of the pore followed by a series of acid residues heading inwards (Glu-159, Gln-170, Asn-167 and Glu-171). The three Glu-171 residues on the inner pore also serves as an iron nucleation site and whilst it is not ruled out that this channel serves as an iron translocation

route, the unusual presence of iron at the "Dps-like" pore suggests that this pore is actually an alternative iron passage (Cuypers et al., 2007). The presence of positive residues on the external surface of the "Ferritin-like" pore is not uncommon. It is thought to create an electrostatic guide for movement of cations, directing them toward the negative pore (Carrondo, 2003). The positive charge is unlikely to inhibit movement of positive metals through the pore as evidenced by the intra-pore localisation of zinc and cadmium ions in the metal soaked crystals of *H. pylori* Dps (*HP*-NAP – PDB ID 4EVB). The *HP*-NAP has three Lys-115 residues that sit at the external opening of the pore which obviously do not impede the translocation of metals (Yokoyama et al., 2012). In other proteins where the pore is lined with fewer acidic residues, like the two *Lactococcus lactis* Dps which only have two (Glu-139 and Glu-140) as opposed to the four of *L. innocua* and *E. coli* Dps, it is proposed that iron oxidation under aerobic conditions is significantly reduced as a result of insufficient iron transportation via the pores (Stillman et al., 2005). This highlights the relationship of these structures to ferroxidase activity and shows how the stabilisation of these regions is integral to maintaining functional activity.

The interface is maintained by both hydrophobic and hydrophilic interactions that help to mediate not only the stability of the pore but overall dodecameric assembly. The quantities of bonds and buried surface areas differ between Dps. The two *Mycobacterium smegmatis* Dps (*Ms*Dps1 – PDB ID 1VEI and *Ms*Dps2 – 2Z90) differ in their surface area and hydrogen bonding (Hbs) across the interface; *Ms*Dps1 has 1601 $Å^2$ buried and 5 Hbs while *Ms*Dps2 has 1987 $Å^2$ and 13 Hbs (Roy et al., 2008). It is proposed that higher number of salt bridges at this interface, in addition to other interfaces within the dodecamer, supply an elevated stability to the *Thermosynechococcus elongatus* Dps (*Te*Dps – PDB ID 2VXX) in high temperatures that is otherwise absent at room temperature when compared to mesophilic Dps (Franceschini et al., 2006). This trait is also displayed by the halophilic *H. salinarum* Dps. On the other hand, prediction and analysis of the interfaces by the online Proteins, Interfaces, Structures and Assemblies server (PISA) of the *D. radiodurans* Dps1 suggests that its N-terminal trimer interface is the only one of the interfaces to not supply a significant contribution to the assembly, this being abnormal among Dps. In spite of the 7 Hbs across the interface, co-ordinated metals throughout the structure, including metals within this "ferritin-like" channel provide a larger complex formation significance score (CSS) (Reon et al., 2012) which places an emphasis on the requirement of metals to increase the stability of the complex.

The importance of residues that contribute to protein self-assembly and stability were explored using the PISA analysis and a further computational "virtual alanine scanning" method in order to select energetically important residues for *in vitro* mutational studies. Two residues (Arg-83 and Arg-133) at the "ferritin-like" interface of *Ec*Dps were predicted to be the most energetically significant residues. This was verified when their mutagenesis to alanine gave rise

to the least stable dodecamers within the study, also resulting in a pronounced abundance of non-dodecameric dimers. The double mutation of these residues culminated in the abolishment of dodecameric assembly and highlighted the importance of these residues not just to their roles in iron transport for example, but also in stabilising the complex (Zhang et al., 2011). In contrast, the stability of *L. innocua* Dps and its natural mutant *L. monocytogenes* Dps (*Lm*Dps – PDB ID 2IY4) were analysed and found the mutations of *Li*Dps Lys-114 to Gln and Asp-126 to Asn that create the *Lm*Dps provide an assembly that is more thermostable at low pH than the *Li*Dps (Bellapadrona et al., 2007).

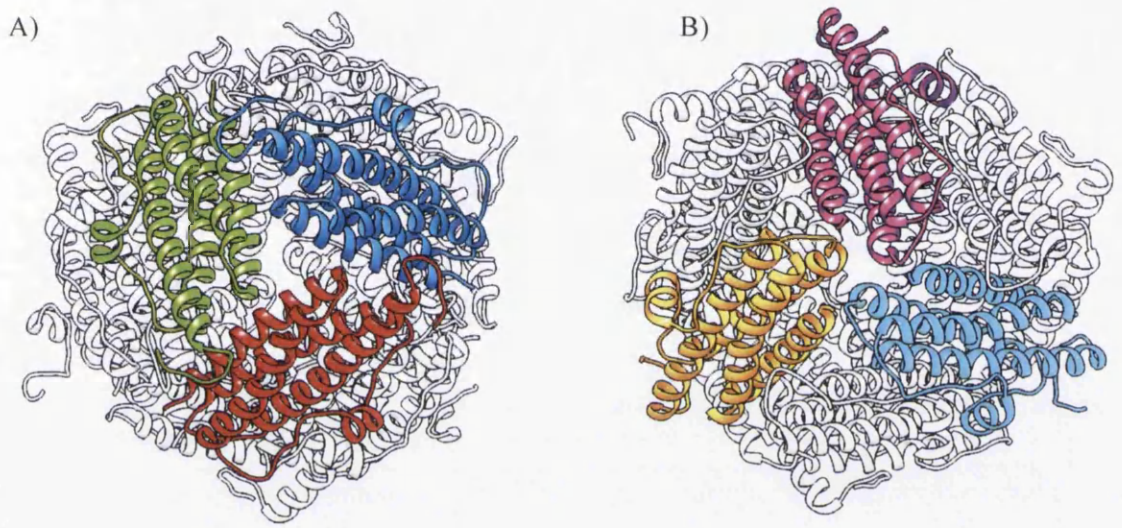


Figure 1.3: The three-fold symmetry interfaces present in the *E. coli* Dps (*Ec*Dps, PDB: 1DPS). The "ferritin-like" three-fold axis (A), in most Dps this structure creates a channel through the protein cage to the internal cavity. The "Dps-like" three-fold axis is not well conserved in sequence (B). Images were created using UCFS Chimera 1.7rc.

1.4.4 "Dps-like" Trimeric Interface – Divergence in Structure and Function

The "Dps-like" trimer (Figure 1.3B), composed of the C-terminal regions of subunits arranged with three-fold symmetry shows a greater variation in sequence and function among Dps. In contrast to the N-terminal pore described in the previous section, the C-terminal pore often displays at the very least a partially hydrophobic pore. The pore itself can also adopt differing contours between Dps; in some cases the channel is blocked by side chains; Phe-34 of *HP*-NAP (Zanotti et al., 2002), Thr-37 and Gln-34 of Dlp1 and Thr-39 and His-36 of Dlp2 from *B. anthracis* (*Ba*Dlp2 – PDB ID 1JIG) (Papinutto et al., 2002). These blockages make it unlikely for the channel to be used as a passage for larger atoms. In contrast to the pattern of negative, hydrophilic N-terminal pores and hydrophobic C-terminal pores, the Dps-like protein

of *S. solfataricus* displays a C-terminal that is composed of acidic residues and despite a constricted region at the centre of the channel, it still retains a large diameter of 4.7 Å at this point (Gauss et al., 2006). This would still provide a more than adequate space required for the movement of iron, which in its ferrous form possess an ionic radii of ~ 0.61 Å (Shannon, 1976). The *D. radiodurans* Dps has a C-terminal channel that is defined by a series of layers of differing groups of amino acids. The channel starts with two layers of positive residues (Arg-205 and Arg-89) followed by three symmetry-related Phe-90 creating a hydrophobic layer. The cavity end of the channel is composed of negatively charged Asp-93 which co-ordinate iron. It is therefore suggested that this negative region could provide a means to direct iron outward, and thus afford an iron exit route (Romão et al., 2006). An alternative hypothesis, given that a metal always occupies this Asp site, is that it functions to stabilise the assembly.

As documented in the aforementioned rational disruption of *Ec*Dps assembly by "virtual alanine scanning", the surface area buried at the *Ec*Dps "Dps-like" interface is low and due to its hydrophobic nature contains no electrostatic interactions, therefore does not contain any hot spots. A similar point was addressed when comparing Dps from mesophilic bacteria such as *E. coli, L. innocua* and *M. smegmatis* with extremophilic Dps from *T. elongatus* and *H. salinarum* (Franceschini et al., 2006). A distinct increase in hydrophilic interactions, particularly salt bridges, were found at the "Dps-like" pores of the extremophilic Dps. When combined with the increase in buried surface area in these Dps, it provides evidence for this region providing extra stability for the dodecamer. Given the sequence variability in these regions it is possible to suggest that in extremophile Dps, the "Dps-like" pore has been adapted and selected to provide strength for the dodecameric assembly.

1.4.5 Dimer Interface and the Ferroxidase Centre

The dimer interface is composed of two anti-parallel subunits adjoining each other so that the two sets of A and B helices form an inter subunit four helix bundle along the extended axis of the monomer (Figure 1.4). The two BC helices, also contributing to the interface, are eventually positioned to the external surface of the dodecamer, creating a two-fold axis of symmetry. The dimer is vitally important to the function of the protein. It not only creates the inter-subunit ferroxidase centres, but its extended buried surface area upon formation and inter subunit hydrophilic interactions to help stabilise the dimer, providing the initial assembly step in the pathway that leads from monomer to dodecamer (Grant et al., 1998). Twelve ferroxidase centres are generated in the dodecamer; two between each dimer with the residues that compose the FOC supplying most of the highly conserved residues within the Dps family. Because of the anti-parallel nature of the dimer, one FOC is located toward one end of the dimer, and the second FOC toward the other with a distance of ~24 Å in between (Ren et al., 2003). This

positions the FOC close to a "ferritin-like" pore, thus each pore has three fairly close FOCs. Despite the fact that in Ftn and Dps the N-terminal three-fold pores supplies a route for iron translocation, the position of the catalytic site is different. The FOCs are located in a shallow groove between the A and B helices of each subunit with each of these helices also supplying the residues involved in iron binding and iron oxidation.

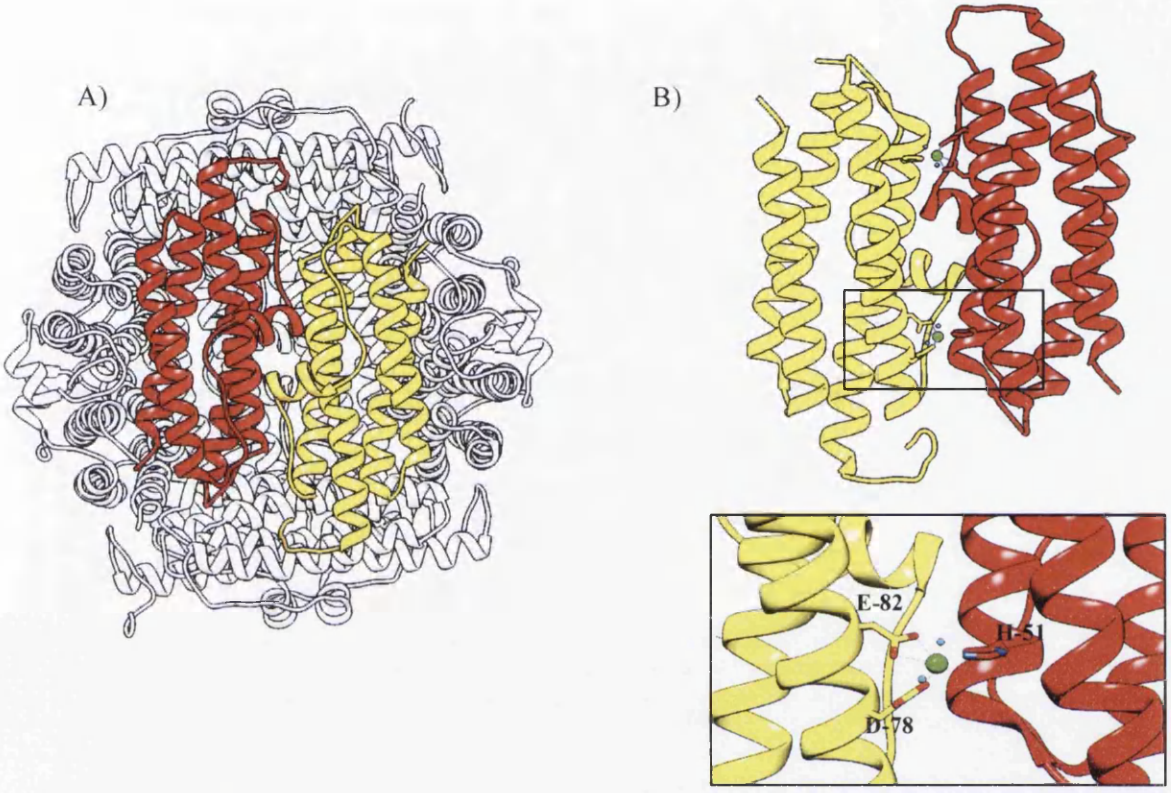


Figure 1.4: A dimer composed of two antiparallel subunits is shown within the dodecamer of *Ec*Dps (PDB ID 1DPS). The dimer is coloured red and yellow (A). The dimer interface creates an environment for the formation of an inter subunit ferroxidase centre. The *Ec*Dps FOC (PDB ID 1F30) (B) is displayed coordinating a zinc atom (green sphere) using residues His-51, Asp-78 and Glu-82 in addition with two water molecules (cyan spheres). Images were created using UCFS Chimera 1.7rc.

Because the FOC is located on the inner surface of the protein, and not within the monomer bundle as for Ftn, it is suggested that the cage plays no conserved role in the mineral nucleation and growth as it does in ferritin L-chain for example (Schwartz et al., 2010), although various negative pockets of residues can adopt this role in a protein specific manner. The oxidation process also differs from Ftn. Unlike ferritin, the Dps have been shown to utilise hydrogen peroxide as a more efficient oxidant for the iron oxidation process with rates of

oxidation 100-fold more effective than with O_2 (Su et al., 2005); (Chiancone and Ceci, 2010). The ability to use H_2O_2 is greatly associated with pathogenic bacteria survival; conferring protection from host defence mechanisms. This survival ability has been demonstrated in bacteria such as *S. enterica* (Halsey et al., 2004) and *C. jejuni* (Ishikawa et al., 2003), where mutation of the corresponding *dps* genes leads to enhanced sensitivity to oxidative stress.

The amino acids that compose the FOC are conserved histidines and carboxylate residues. There are two exceptions to the Dps FOC rule that have been found, as neither of the two *L. lactis* Dps proteins have conserved ferroxidase centres. Instead, the presence of a novel intra-subunit metal binding site was detected at the base of its novel N-terminal tail helix. At this site His-22 and His-33 bind a metal using two water molecules (Stillman et al., 2005). Additionally, the Dps-like protein of *S. solfataricus* possesses a Bfr-like FOC which is buried within the four helix bundle (Gauss et al., 2006). The *Li*Dps complexes iron using Asp-58 and Glu-62 from one subunit and His-31 from the symmetry-related partner subunit. A water molecule is also utilised, hydrogen bonded to His-43 (Ilari et al., 2000). This structure would suggest the presence of a mono-iron centre only, however the di-iron centre was modelled, replacing the water with iron and allowing the Glu-62 to supply a carboxylate bridge between irons and also links in the water molecule occupying the site prior to the addition of the second iron. This replicates the A and B sites found within Ftn proteins. Subsequent work carried out to investigate the importance of the His residues to *Li*Dps found they were essential for the formation of a hydrogen bond network that maintained the co-ordination of iron and water molecules, confirming that a di-iron centre is probable (Ilari et al., 2005, Gauss et al., 2006). It also demonstrated the importance of the His residues to the conformations of the Glu and Asp residues associated with the FOC and suggested that the iron binding sites are pre-formed and do not rely on side chain conformational changes in order to accept iron. The di-iron site of the *B. brevis* Dps (*Bb*Dps – PDB ID 1N1Q) was revealed to be highly similar to *Ec*Dps and *Li*Dps. However, rather than modelling the di-iron site, direct evidence was used to prove its existence. The presence of two μ-oxo bridges were predicted by difference electron density maps to occur between the two iron atoms, and between His-43 and the B site iron (Ren et al., 2003).

The examination of the *D. radiodurans Dr*Dps2 FOC questions the function of the conserved Trp-71 residue that would appear to modulate the ability of FOC residues to chelate iron. Once the Trp-71 hydrophobic side chain has flipped 180° out of the iron binding sites, the salt bridges stabilising Asp-97 and Glu-101 side chains are disrupted, causing conformational changes. These changes amount to a positional change of 1.7 Å for Asp-79 and 3.7 Å for Glu-101 between Cα atoms (Cuypers et al., 2007). The organisation of the FOC differs in this species when compared to other Dps in spite of the fact that the residues are largely conserved between the Dps, however it is suggested that the variances could be attributed to the stages of

iron mineralisation at which the structures were observed, yet leaves the exact role of the Trp residue open for further insight. The *HP*-NAP FOC Trp-26 also switches between two rotameric conformations of 180° and is suggested to regulate ferroxidase activity (Tsuruta et al., 2012). Other studies have revealed that the conserved Trp in combination with another conserved Tyr residue are likely to serve as hydroxyl traps. Both residues are located close to the FOC and prevent the release of intra protein generated hydroxide radicals formed by the transfer of electrons which are generated during the oxidation process (Bellapadrona et al., 2010). If this role was to be correct, it implicates a use for the protein cage itself in protection of DNA and macromolecules and not just as a scaffold in which to store the iron.

The differences in mechanics of the ferroxidase sites were also explored for the two *B. anthracis* Dps (*Ba*Dlp1 – PDB ID 1JI5 and *Ba*Dlp2). Despite sharing ~60% sequence homology (Liu et al., 2006), the two proteins displayed different iron binding affinities with Dlp2 binding iron 30 times stronger than Dlp1 at the first iron binding site. In both cases the binding of a second iron is two to three orders in magnitude weaker than the first iron (Schwartz et al., 2010).

A difference in coordination capacity at the FOC was proposed to be the main difference between the iron (III) translocation process in *Li*Dps and *Ec*Dps. In *Ec*Dps, a salt bridge between Lys-48 and Asp-78 (an iron binding residue) is suggested to decrease the affinity for bound Fe (III), which aids its movement after oxidation to the cavity. This is the reason why the crystal structures of iron loaded *Ec*Dps exhibits two water molecules at the FOC while the absence of a salt bridge in *Li*Dps means iron atoms are held at the FOC and are detected in the structures (Ilari et al., 2002).

The surface area buried upon complexation is usually the most extended of all the interfaces present within Dps. In all cases documented the interface is always above 1000 $Å^2$/monomer. However, as for the other interfaces discussed, Dps from extreme environments tend to support an extended buried surface area with an increased abundance of hydrogen bonds and salt bridges which are said to confer enhanced stability (Franceschini et al., 2006). The stability of the dimer has been demonstrated in a range of conditions both for Dps and Ftn. Denaturing environments used to test oligomeric assembly and stability usually result in dissociation of the dodecamer down to a dimeric state (Chiaraluce et al., 2000); (Yoshizawa et al., 2007). The failure to disrupt the assembly any further or the requirement of enhanced extreme conditions to break apart the dimer emphasises the importance of this structure to oligomeric assembly and stability.

1.4.6 Differential Roles for the Tails of Dps – Mediating DNA Binding, Oligomeric Assembly/Stability and Virulence

The Dps "tail" structures protrude from N- and C- terminal regions of the four helix bundles and in most cases are placed to the exterior of the assembled cage with the exception of the *S. solfataricus* Dps-like protein in which the C-terminal tail actually projects into the interior of the protein (Gauss et al., 2006). The exact sequence, length and position of tails show poor conservation; nevertheless they still display some functional conservation. The two main features of Dps in which the tails are implicated are oligomeric assembly/stability and non-specific DNA binding ability. Furthermore most of the previously mentioned diverse functions of the Dps are manifested by slight structural deviations of the Dps monomer, particularly at the N-terminal tail.

Even though the exact structural factors of the Dps mediated DNA interaction are still unclear, it is strongly linked with an extended N-terminal tail region, which displays multiple positively charged residues. The *Ec*Dps contain an N-terminal tail of 22 residues long. The flexibility of this tail is emphasised by the ability to only resolve the latter 9 residues, leaving the first 13 residues disordered in the structure (Grant et al., 1998). In spite of this, deletion mutants where the first 8 and 18 residues were removed displayed an altered ability to bind and condense DNA. However, no impact on oligomeric assembly was detected (Ceci et al., 2004). Emphasis has been placed on the lysine rich regions of the N-tail to mediate the interaction. The ability to bind and condense the DNA provides the *Ec*Dps with a dual-mode protection; firstly, binding, condensing and physically shielding the DNA from damage such as DNase mediated cleavage, and secondly by ferroxidase activity protecting against oxidative damage. When interacting with DNA, the *Ec*Dps is found to rapidly co-crystallise to produce highly stable structures both *in vivo* and *in vitro*. The hexagonal lattice structures of Dps/DNA crystals indicate that DNA is shielded in layers between Dps where the dodecamers are all aligned in the same fashion (Wolf et al., 1999). This was further reinforced with the multilayer crystal structure of *Bb*Dps. This structure, in the absence of DNA, suggested that the DNA could be located in grooves between hexagonal layers of Dps and that slight sliding between layers could even lead to the accommodation of supercoiled DNA. This would result in a DNA/Dps structure that still allows the Dps to perform their iron scavenging roles (Ren et al., 2003). However whilst it was noted that *Bb*Dps is capable of binding DNA, evidenced by band shifts in DNA electrophoresis, there is an absence of a long N-terminal tail and no C-terminal tail, thus DNA interaction in this case is not mediated by the tails.

The two *L. lactis* Dps (*Ll*DpsA – PDB ID 1ZUJ and *Ll*DpsB – PDB ID 1ZS3) are highly similar; both contain an unusual surface exposed helical structure present in the N-terminal tails. These structures are perpendicular to the core bundle and eventually extend across an adjacent symmetry related monomer with a helix at each end of the dimer (Stillman et al., 2005). This helix has been implicated in DNA binding ability, which results in amorphous Dps/DNA complexes. Deletion of the helix impaired DNA binding ability. Since the *Ec*Dps DNA binding is associated with positively charged Lys residues within its N-terminal tail, three N-tail helix Lys residues in *Ll*DpsA were mutated, however this did not disturb the ability of the protein to interact with DNA. On the other hand, metal chelation through the addition of EDTA did inhibit DNA interaction signifying a role for cations. The loss of the N-tail helix did not affect dodecameric assembly thus this tail has no role in assembly, however dodecameric stability was not investigated. There are other examples of N-tail helices. The *K. radiotolerans* Dps (*Kr*Dps – PDB ID 4A25) has a small N-helix which mediates hydrophobic interactions with adjacent subunits within the dodecameric structure (Ardini et al., 2013). Nevertheless the significance that these interactions have for assembly remains unknown; furthermore the DNA binding ability of *Kr*Dps has yet to be investigated.

The *D. radiodurans Dr*Dps1 presents a short coil within its very long N-terminal tail (55 residues), and once assembled the coil is located on the external surface of the dodecamer. The tail is essential for DNA binding ability and is important for oligomeric assembly (Bhattacharyya and Grove, 2007). The N-helix mediates interactions with the major groove of DNA thus allowing DNA to be bound along each axis of the six dimers that compose the dodecamer. This provides a model for the hexagonal packing of DNA and Dps layers where each dodecamer interacts with the DNA layer above and below (Nguyen et al., 2012). The contribution that the N-tail gives to assembly and stability is intricate. A novel metal binding site located in a loop between the N-helix and the A helix of the core bundle is reported to serve as an assembly switch. Upon binding of cations to these sites the dodecameric assembly occurs (Grove and Wilkinson, 2005). Deletion of 33 residues of the N-tail up to the metal site does not affect dodecameric assembly; however deletion of the full 55-residue tail not only limits assembly to a dimeric state, but also eliminates DNA interaction. The use of metal chelators did not influence the ability of the protein to interact with DNA, thus it is suggested that electrostatic interactions between the numerous positive residues within the N-terminal tail and the negative phosphate backbone of DNA are the basis of *Dr*Dps1 DNA interaction.

The *Dr*Dps2 protein has an equally unusual C-terminal tail. This tail also contains a small helix, which is also accompanied by a novel iron-binding site. The tail region was predicted by PISA analysis to be important to preserving the C-terminal "Dps-like" trimer, which in this protein maintains an abnormally high CSS when compared to other Dps. In a C-

tail deleted mutant of *Dr*Dps2, assembly was completely absent, presenting the protein entirely as a monomer, a phenotype rarely seen in Dps mutants. Furthermore it was hypothesised that the metal present at the binding site maintained a tail conformation that was compatible with DNA binding. Loss of metal meant loss of DNA interaction (Reon et al., 2012).

M. smegmatis contains two Dps (*Ms*Dps1 and *Ms*Dps1). *Ms*Dps1 is characterised by a C-terminal tail that is involved in both DNA binding and oligomeric assembly (Gupta and Chatterji, 2003, Ceci et al., 2005). The 26 residue C-terminal tail can be split into a bi-functional entity. Removal of the last 16 residues abolishes DNA binding but maintains dodecameric assembly, while complete removal of the tail leads to an unusual open decameric structure (Roy et al., 2007). The N-terminal tail was also implicated in oligomeric assembly where Van der Waal forces link the first four residues of the N-tail with 5 residues of the C-terminal tail at the interface of the N-terminal "ferritin-like" pore. Loss of the N-terminal tail disrupts assembly and also interrupts DNA interaction, although it is thought that this is an indirect effect of the C-terminal tail losing its configuration. The *Ms*Dps2 possesses a longer N-terminal tail that is also involved in bracing the dodecamer as the C-terminal tail of *Ms*Dps1 does. Present on the surface of the dodecamer, the N-tail not only braces an adjacent subunit, but goes further on to interact with the adjacent subunit's dimer partner where Arg-4 of the N-tail interacts with Asp-86 from the BC helix of the furthest subunit (Roy et al., 2008). The N-tail has been postulated to interact with DNA and unlike *Ms*Dps1, promotes nucleoid condensation *in vivo* (Saraswathi et al., 2009).

1.4.7 Virulence in Dps

Whilst DNA interaction and oligomeric assembly/stability have been described as functions of the N and C terminal tails, there are other roles adopted by the tails of Dps. The *T. pallidum* Dps (*Tp*F1 – PDB ID 2FJC) possess immunogenic properties, which originate from its flexible N-tail. However, because of the plasticity in the tail, the explanation as to how the tail fulfils this function remains undefined (Thumiger et al., 2006). The *HP*-NAP from *H. pylori* has the ability of binding and condensing DNA. However it is not a tail that affords this capability. Instead it is proposed that protonation of the cage's exterior amino acids creates a positively charged protein surface, uncharacteristic of Dps and this mediates the interaction with DNA (Ceci et al., 2007). Furthermore this surface is the origin of the protein as a major virulence factor for *H. pylori*, hence the name *H. pylori* neutrophil-activating protein (D'Elios et al., 2007). Once secreted, it acts as pro-inflammatory molecule, inducing inflammation which is associated with a release of nutrients from inflamed tissue; consequently promoting *H. pylori* growth (de Bernard and D'Elios, 2010).

1.4.8 Dps and Biotechnology

The extensive research into ferritin illustrated its potential as a platform to develop a broad range of biotechnological applications. The development relies on their interesting biochemical features; a strong protein cage and the ability to create size constrained iron oxide nanoparticles. Of significant interest is the ability to mineralise and deposit metals within the cavity. This process has been manipulated by means of reaction condition control and protein sequence/structure alteration for the outcome of artificial synthesis of metal complex nanoparticles (Yamashita et al., 2010). The iron core that can be deposited within the cavity of the Ftn is usually cubic, polycrystalline ferrihydrite (Pan et al., 2009). In this form the core displays paramagnetic behaviour, the nature of which is dependent on the protein used and iron content of the core (fully loaded, partially loaded) in addition to the method employed to establish the core (Michel et al., 2010). Through systematic mutational studies of proteins and their cavities, not only has valuable data regarding their biological function been generated, but also methods of manipulating core formation have evolved (Haikarainen et al., 2011). Furthermore metals such as nickel and chromium (Okuda et al., 2003), copper (Ensign et al., 2004), and even gold (Sun et al., 2011) have also been deposited within the cavities and bestow alternative properties to the nanoparticles, expanding their potential.

Cage assembly and modification has also been explored. Using the crystal structures of proteins which have initially been investigated to explore their biological functions, computational approaches have led to cages which display enhanced self-assembling properties and increased thermal stability (Ardejani et al., 2013). The additions of functional groups on the surface of the proteins also provide a means of equipping the protein with a desired purpose, often described as "active targeting". The addition of ligands to the surface of the protein cages can be carried out in multiple ways. Carbohydrates such as N-acetyl-D-glucosamine were added to a ferritin surface via glycosylation with amine surface groups (Valero et al., 2011). Chemical modification in a spatially controlled manner has also been achieved using solid surface supports and a masking/unmasking approach to polarise the C-termini of subunits within the cage (Kang et al., 2009). These sorts of cage modifications together with the mineralised cores produce proteins with a range of biotechnological uses that have been tested both *in vitro* and *in vivo*. These technologies include: cellular targeting and labelling (Li et al., 2011), drug delivery (Zhen et al., 2013), MRI contrast agents (Geninatti Crich et al., 2012), bio-separation (Haun et al., 2010) and magnetofection (external magnetic targeting) to name just a few. However the exploration of Dps is less advanced even though they offer different benefits: diversity in size and chemical attributes. Furthermore, the discovery of novel Dps from harsh environments can also provide platforms, which offer distinct biochemical properties such as extreme cage stability.

Another avenue for the use of Dps involves the immuno capabilities of proteins such as HP-NAP. The immune system's Th1/Th2 balance is thought to be important in immunopathology response and control. Polarisation of either Th1 dominance or Th2 dominance is thought to lead to various diseases and infections (Romagnani, 2003, Matsuda et al., 2007). However certain circumstances exist when a Th1 polarised response would be beneficial, for example cancer therapy, or in the restoration of the Th1/Th2 balance after polarisation of a Th2 response. Therefore the ability of the HP-NAP protein to coordinate a Th1-polarised immune response as part of *H. pylori's* host infection has led to research into utilising this function as a tool for immunotherapeutic purposes (D'Elios et al., 2007). There are examples of HP-NAP exhibiting positive results in *in vitro/in vivo* studies such as decreasing tumour growth of invasive bladder cancer (D'Agostino et al., 2012). In addition, it has been demonstrated that HP-NAP can be delivered by a tissue specific adenovirus. Once targeted infection to a range of tumour cell lines was achieved, intracellularly released HP-NAP promoted micro-environmental immune response, causing Th1 immune polarisation thus increasing the oncolytic effects (Ramachandran et al., 2013).

1.5 *Streptomyces coelicolor* – Differential roles of the three Dps

Streptomycetes represent the largest genus of the actinobacteria and are represented by the well-studied and genetically characterised *Streptomyces coelicolor*. The genome of this soil dwelling, filamentous bacteria was published in 2002 (Bentley et al., 2002) and presented a high GC content (72.1%) synonymous with other streptomycetes (Ou et al., 2003). The high number of predicted protein encoding genes (7825), considered at the time to be one of the largest for any bacterium, is a representation of the complexity of the organism. The sequenced genome was important for the study of streptomycetes, as they are considered economically and medically significant due to the plethora of bioactive secondary metabolites, which they produce. These metabolites have been harnessed for uses that range from antibiotics to anti-tumour agents and insecticides (Steingrube et al., 1997).

The production and secretion of secondary metabolites would appear to have co-evolved with the development of a complex multicellular life cycle as a consequence of the competition for nutrients with other microorganisms (Chater, 2006). The life cycle is split into three distinct stages: filamentous vegetative growth is followed by the development of aerial hyphae and culminates in the septation of the aerial hyphae for the formation of unigenomic spores (Paradkar et al., 2003). The cycle is under the control of numerous genes; the genome of *S. coelicolor* was demonstrated to contain unprecedented quantities of regulatory genes (Bentley et al., 2002).

Mycelial growth is initiated from spore germination, which occurs under favourable conditions, usually in response to the availability of soluble nutrients. Vegetative growth takes places via apical hyphal extension (Gray et al., 1990), forming a network of densely packed, branched filaments composed of compartments often housing multiple nucleoids (Elliot et al., 2003). During this phase of the cycle *Streptomyces* secrete extracellular hydrolytic enzymes that facilitate the breakdown of the nutritional resources, particularly insoluble matter. The exportation of these enzymes precedes the production and secretion of secondary metabolites which are produced from primary metabolite intermediates (Champness, 1988). This strategy has evolved as a protective mechanism, inhibiting other microorganisms from hijacking the solubilised material generated as a nutrient source.

Upon detection of nutrient depletion, morphological differentiation occurs, where aerial branches develop, sending hyphae to "reach for the sky" (Claessen et al., 2006). In parallel to the progression of aerial hyphal development, programmed primary hyphal death occurs in a controlled and stepwise manner and the re-generated nutrients are channelled upward to facilitate the growth and maturation of the aerial hyphae. The life cycle is concluded with the process of spore formation. A series of genetic and physiological processes occur within compartments that see a high degree of synchronous chromosomal replication (Ruban-Osmialowska et al., 2006) followed by chromosomal condensation and segregation to form the pre-spore compartments in between the regularly spaced sporulation septa (Flardh et al., 1999). These compartments eventually mature to from individual and separated spores, each containing a single chromosome. In *S. coelicolor*, a coupling exists between developmental control and stress response regulation of certain protein groups. Referred to as "synexpression groups", it was found that the separately controlled stress response genes were also modulated developmentally, potentially allowing the cell to adapt to the various stresses it encounters as a result of the developmental program or even as control mechanisms for development itself (Vohradsky et al., 2000).

The three Dps of *S. coelicolor* have been shown to have dual regulation in response to stress and during development. The expression of the Dps proteins: *Sc*DpsA, *Sc*DpsB and *Sc*DpsC are induced in response to stress in a stimulus-dependent manner where osmotic shock (250 mM KCl), heat shock (42°C) and nutrient shock (no carbon source) all induce expression of at least one Dps (Facey et al., 2009). The expression of the *Sc*Dps proteins also exhibits developmental regulation whereby expression of *Sc*DpsA gradually increases through the developmental cycle to a maximum concentration within the spores. Furthermore, disruption of the *dpsA* (SCO0596), *dpsB* (SCO5756) and *dpsC* (SCO1050) genes results in significant morphological differences and irregularities in nucleoid condensation during the development of spores. This indicated a role for the proteins in DNA condensation and is the first example of

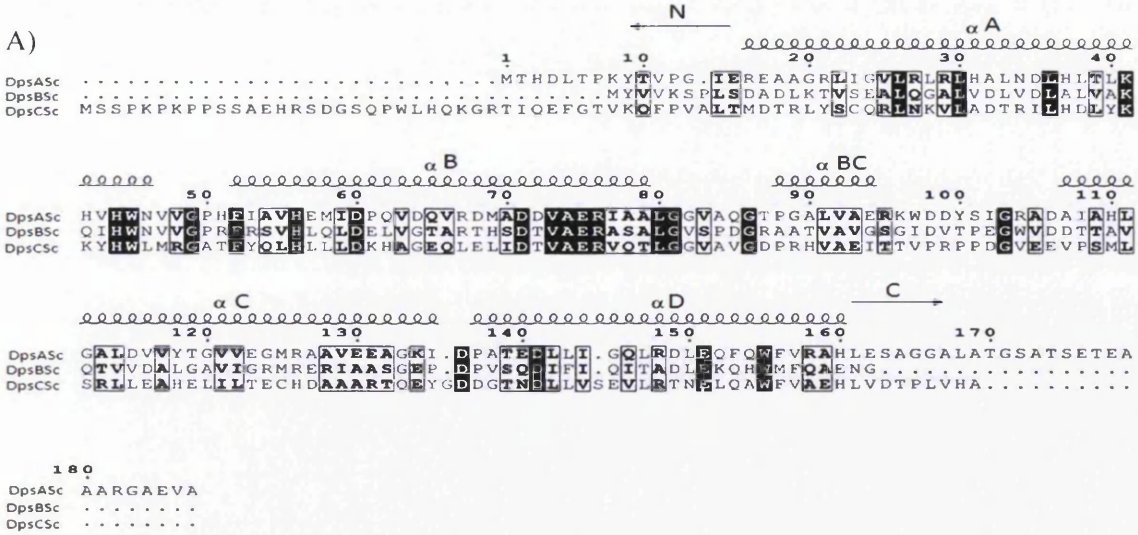
Dps playing a role in bacterial cell division. However, the Dps were not implicated in the oxidative stress response, nor were the mutant strains sensitive to oxidative stress.

The investigation of *Sc*DpsA expression regulation revealed a complex network of sigma factors that control and modulate expression from a single promoter (Facey et al., 2011). SigB-like sigma factors are associated with the expression of *dpsA*, providing *in vivo* evidence of "cross-talk" in promoter specificities. SigB and SigH sigma factors are implicated in direct modulation of transcription from the *dpsA* promoter. Furthermore, the *dpsA* promoter was shown to be a direct target for the WhiB sigma factor, the first example of its kind in *S. coelicolor*.

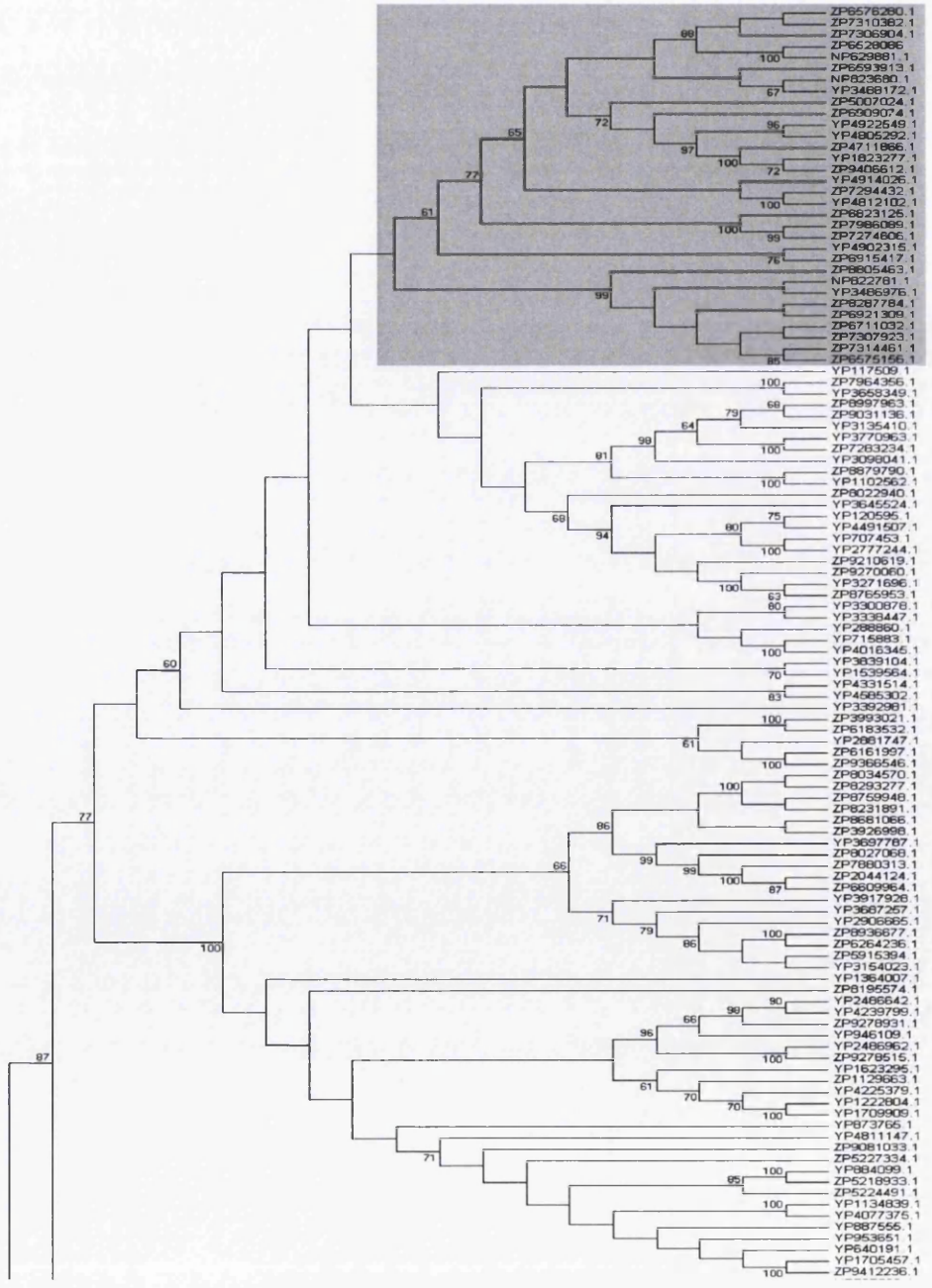
There are a number of peculiarities regarding the *Sc*Dps. It is unusual for an organism to contain three *dps* genes. However, given that the proteins are differentially expressed, it could point to some form of functional divergence. In parallel to the work carried out in this study, phylogenetic analysis pertaining to the evolutionary history of the three *dps* genes within *S. coelicolor* was also undertaken and yielded interesting results. Reconstruction of the phylogeny of actinobacterial Dps sequences remarkably revealed that the three *Sc*Dps proteins are separated into three distinct clades: a DpsA-like clade, DpsB-like clade and DpsC-like clade. This is illustrative of a unique evolutionary history for each of the *Sc*Dps in recent history (Facey et al., 2013). Furthermore, the study also revealed that the DpsC orthologs have a very narrow and rare distribution in bacteria; additionally DpsA orthologs are rare in *Streptomyces*. These two findings point to the proteins having been laterally acquired and reinforce the notion that they are functionally distinct and can only fulfil certain roles within the cell.

Protein sequence alignments of the *Sc*Dps displayed with secondary structure prediction in Facey *et al* (2009) revealed that each of the *Sc*Dps proteins has varying length N and C-terminal tails. Intriguingly the phylogenetic reconstruction mirrored the pattern of *Sc*Dps tail morphology; DpsA like proteins contain an N-terminal tail and a long C-terminal tail, DpsB-like proteins contain very short or no tails and the few DpsC-like proteins contain very long N-terminal tails. Despite the previously mentioned poor sequence conservation observed in the tail regions, it would appear that, at the very least, tail length and position is conserved and thus implies that this feature does have a functional role within the Dps proteins in a protein specific manner

A)



B)



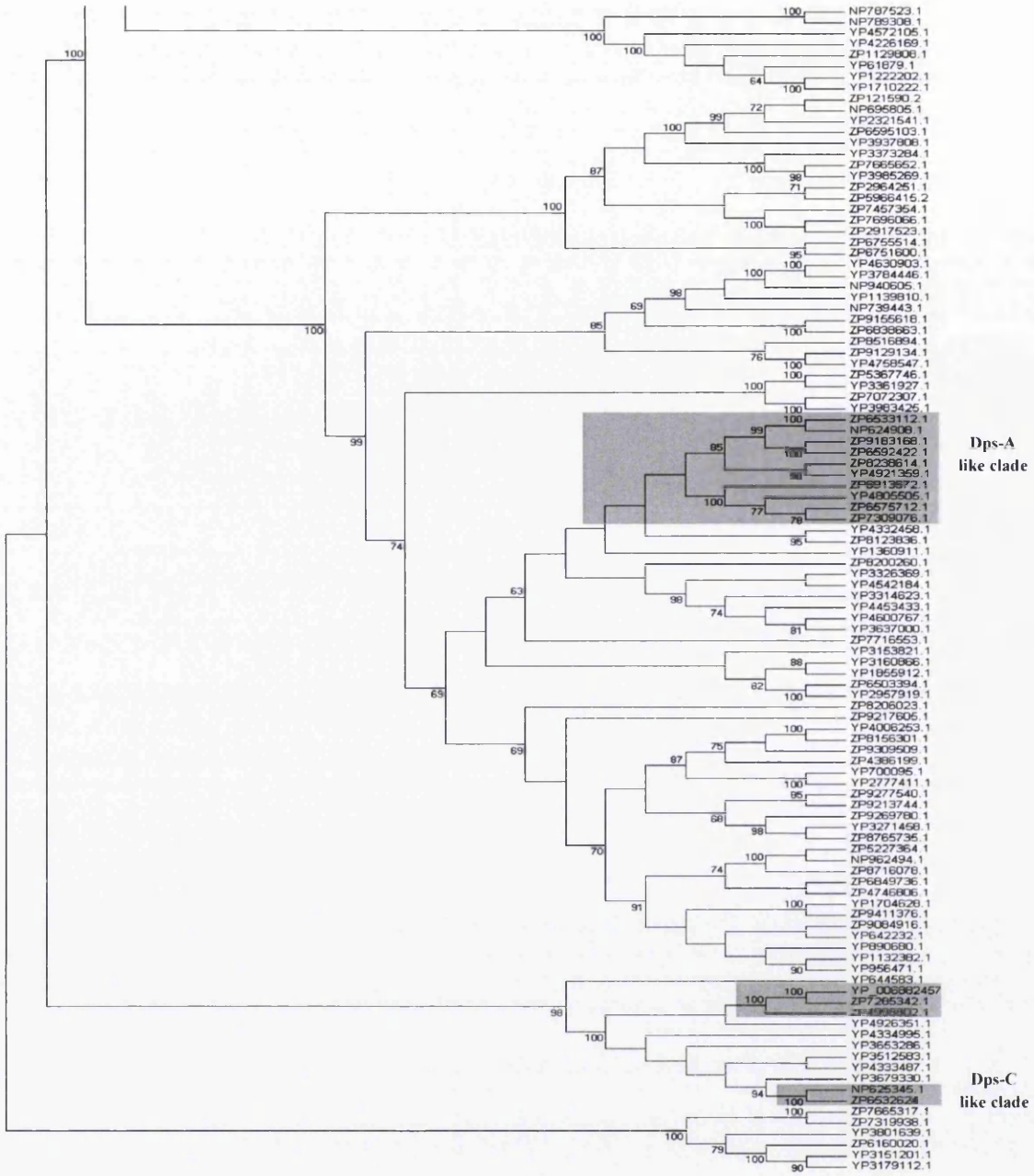


Figure 1.5: Amino acid sequence aligment of ScDpsA, ScDpsB and ScDpsC and reconstruction of the phylogeny of Actinobcterial Dps orthologous proteins. (A) The amino acid sequence alignment shows the characteristic five helices along with the terminal tails. Positions of helicies are demonstrated above the sequence with the helix labels (A, B, BC, C and D). N and C terminal tail positions are highlighted with use of arrows above tail regions. (B) The three ScDps proteins display unique evolutionary history evidenced by their grouping on separate clades among other actinobacterial proteins on a majority rule consensus phylogenetic tree. Bootstrap values are given for each major node. Shaded boxes represent Dps proteins within Streptomyces. Its protein accesstion number represents each protein sequence. The labels of Dps-like clade represent the clade that contains ScDpsA, ScDpsB or ScDpsC. Images reproduced from Facey *et al.*, (2013)

1.6 The aims of this study

The intention of this study is to analyse the Dps proteins of *Streptomyces coelicolor* from a biochemical and structural perspective. This thesis aims to:

- Characterise the biochemical functional attributes of recombinant *Sc*Dps *in vitro*. This includes investigating oligomeric assembly, ferroxidase activity and ability to interact with DNA
- Structural characterisation of the DpsA, DpsB and DpsC proteins
- Assess the importance of the N and C terminal tails for the functional properties of the proteins
- Assess the capacity for functionalising the protein cage with bioactive peptides for the purpose of creating novel nano-biotechnological tools.

Chapter 2

Materials and Methods

2.1 Bacterial Strains

Bacterial strains used in this study are listed in table 2.1 below.

Table 2.1: Bacterial strains

Bacterial Strain	Genotype	Source
E. coli JM109	F′ *tra*D36 proA $^{+}B^{+}$ lacIq Δ (lacZ)M15/ Δ (lac-proAB) glnV44 e14- gyrA96 recA1 relA1endA1 thi hsdR17	Yanisch-Perron *et al.* (1985)
E. coli BL21 (DE3)	F^{-} ompT gal dcm lon $hsdS_B(r_B^{-} m_B^{-})$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	New England Biolabs (NEB)
S. coelicolor A3(2) M145	Prototrophic SCP1-SCP2 Pgl^{+}	Kieser *et al.* (2000)

2.2 Plasmids

Plasmids created or used during this work are listed in table 2.2 below. Maps and construction details can be found in the results chapters relevant to their use.

Table 2.2 Plasmids

Plasmid/Cosmid	Marker genes/Description	Source
pDpsA9	pIJ8600, *tipA-dpsA::His6*	Facey *et al.* (2009)
pDpsB9	pIJ8600, *tipA-dpsB::His6*	Facey *et al.* (2009)
pDpsC9	pIJ8600, *tipA-dpsC::His6*	Facey *et al.* (2009)
pDpsA4	*dps*A in pGEM-T Easy	Facey *et al.* (2009)
pDpsB4	*dps*B in pGEM-T Easy	Facey *et al.* (2009)
pDpsC1	*dpsC* in pGEM-T Easy	Facey *et al.* (2009)
pET-16k	$Kanamycin^{R}$, T7 promoter C-His tag expression vector	Novagen
pET-26b^{+}	$Kanamycin^{R}$, T7 promoter N-His tag expression vector	Novagen
pCR2.1	Amp^{R}, Kan^{R}	
pDpsA13	$Kanamycin^{R}$, T7, *dpsA* N-His tag expression vector	This study
pDpsA14	$Kanamycin^{R}$, T7, *dpsA* C-His tag expression vector	This study

Plasmid/Cosmid	Marker genes/Description	Source
pDpsADCT	Kanamycin^R, T7, *dpsAdelCtail* C-His tag expression vector	This study
pDpsADNT	Kanamycin^R, T7, *dpsAdelNtail* C-His tag expression vector	This study
pDpsADTM	Kanamycin^R, T7, *dpsAdelN/Ctail* C-His tag expression vector	This study
pDpsC14	Kanamycin^R, T7, *dpsC* C-His tag expression vector	This study
pDpsCDNT	Kanamycin^R, T7, *dpsCdelNtail* C-His tag expression vector	This study
pDpsAdelBC	Kanamycin^R, T7, *dpsA deleted BC helix* C-His tag expression vector	This study
pDpsB14	Kanamycin^R, T7, *dpsB* C-His tag expression vector	This study
pDpsAmCh	Kanamycin^R, T7, *dpsA-mCherry* C-His tag expression vector	This study
pDpsACD27L	Kanamycin^R, T7, *dpsA-C. diff lysin* C-His tag expression vector	This study
pDpsACD27Lt	Kanamycin^R, T7, *dpsA-C. diff lysin truncated* C-His tag expression vector	This study
pDpsAMR11	KanamycinR, T7, dpsA-MR11 lysin C-His tag expression vector	This study
pDpsA-MR11t	KanamycinR, T7, dpsA-MR11 lysin truncated C-His tag expression vector	This study

2.3 Chemical Reagents

Chemical solutions and buffers were made using distilled water (dH_2O) from a reverse osmosis water purification system by MILLI-RO. Ultra-pure distilled water (ddH_2O) was also provided by a MILLI-Q water purification system and used when required. Solutions were prepared and pH measurements were taken at room temperature. Solutions requiring sterilisation were autoclaved at 121° C. 15 psi for 15 minutes, when this method was not suitable, solutions were filter sterilised using 0.2 μm Millipore filters and syringe. The components of the commonly used buffers are given in table 2.3

Table 2.3: Reagents and buffers

Reagent/Buffer Solution	**Composition**	**Quantity per litre dH_2O unless otherwise stated***
10% Glycerol (v/v)	100% Glycerol	100 ml
20% Glycerol (v/v)	100% Glycerol	200 ml
10X Tris-Borate/ EDTA (TBE)	Tris Boric Acid EDTA Adjust to pH 8.0 with HCl	108.00 g 55.00 g 9.30 g
1X TBE	10X TBE	100 ml
0.5X TBE	10X TBE	50 ml
Bromophenol Blue DNA loading dye (6X)*	Sucrose Bromphenol Blue 1X TBE dH2O	40.00 g 60.00 mg 10.00 ml 90.00 ml
10 % SDS*	Sodium dodecyl sulphate dH_2O	10.00 g Up to 100 ml
1% Agarose*	Agarose 0.8X TBE	2.00 g Up to 200.00 ml
10% Ammonium persulphate*	Ammonium persulphate dH_2O	100.00 mg Up to 1.00 ml
SDS PAGE Running Buffer	Tris Glycine SDS	6.00 g 14.40 g 1.00 g

Reagent/Buffer Solution	Composition	Quantity per litre dH_2O unless otherwise stated*
4X SDS Protein Loading Buffer*	1 M Tris/HCl pH 6.8	200.00 µl
	10% SDS (w/v)	400.00 µl
	100% Glycerol	400.00 µl
	DTT	60.80 mg
	Bromophenol Blue	20 mg
Stacking gel (2X)*	Acrylamide/Bis Acrylamide	650 µl
	dH_2O	3 ml
	1 M Tris pH6.8	1.25 ml
	10% SDS	50 µl
	10% TEMED	5 µl
	10% APS	25 µl
Separating gel (15%)*	Acrylamide/Bis Acrylamide	3.75 ml
	dH_2O	1.75 ml
	1.5 M Tris pH8.8	3.75 ml
	10% SDS	75 µl
	10% TEMED	7.5 µl
	10% APS	37.5 µl
Native PAGE gel (7.5%)	Acrylamide/Bis Acrylamide	1.5 ml
	dH_2O	3.38 ml
	1.5 M Tris pH8.8	3 ml
	10% TEMED	15 µl
	10% APS	40 µl
SDS PAGE running buffer	Tris	6 g
	Glycine	14.4 g
	SDS	1 g
Native PAGE sample buffer (6X)*	BisTris	0.418 g
	6N HCl	107 µl
	Glycerol	4 g
	NaCl	0.117 g
	Ponceau S	0.4 mg
	ddH_2O	Up to 10 ml
Native PAGE Cathode buffer	Tris	6 g
	Glycine	14.4 g
	20XCathode Additive buffer (Invitrogen)	25 ml

Reagent/Buffer Solution	**Composition**	**Quantity per litre dH_2O unless otherwise stated***
Native PAGE Anode buffer	Tris Glycine	6 g 14.4 g
Coomassie Stain*	R250 Coomassie Methanol Acetic Acid dH_2O	250.00 mg 90.00 ml 20.00 ml 90.00ml
Coomassie De-stain*	Methanol Acetic Acid dH_2O	90.00 ml 20.00 ml 90.00 ml
Western Transfer Buffer	Tris Glycine 10% SDS (w/v) Methanol	5.82 g 2.93 g 3.75 ml 200.00 ml
10X Tris-Buffered Saline (TBS) pH 7.6	Tris NaCl Asjust pH to 7.6	24.20 g 80.0 g
1X Tris Buffered Saline-Tween (TBS-Tween) pH 7.6	10X TBS Tween 20	100 ml 1 ml
Blocking buffer (5%)*	TBS/ Tween Blotto Blocking powder	100 ml 5 g
Blocking buffer (10%)*	TBS/ Tween Blocking powder	100 ml 10 g
In gel iron staining*	Potassium ferricyanide 50 mM Tris-HCl, 100 mM NaCl pH 7.5	1.64 g 50 ml
Iron staining colour developing solution*	Tri-chloroacitic acid Methanol dH_2O	5 g 5 ml Up to 50 ml
2.5 M $MgCl_2$	$MgCl_2$	238 g
5 M KCl	KCl	372 g

Reagent/Buffer Solution	**Composition**	**Quantity per litre dH_2O unless otherwise stated***
1.5 M Tris/HCl pH 8.8*	Tris ddH_2O Adjust pH to 8.8	36.33 g Up to 200.00 ml
1 M Tris/HCl pH 6.8*	Tris ddH_2O Adjust pH to 6.8	34.22 g Up to 200.00 ml
5 M Tris/HCl pH 8.5*	Tris ddH_2O Adjust to pH 8.5	171.1 g Up to 200.00 ml
5 M NaCl*	NaCl ddH_2O	146.1 g Up to 500.00 ml
2 M Imidazole *	Imidazole ddH_2O	27.22 g Up to 200.00 ml
Binding/Washing Buffer	5 M Tris-HCl 5 M NaCl 2 M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 25 ml Up to 1 L
100 mM Imidazole Elution buffer	5 M Tris-HCl 5 M NaCl 2M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 50 ml Up to 1 L
200 mM Imidazole Elution buffer	5 M Tris-HCl 5 M NaCl 2M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 100 ml Up to 1 L
300 mM Imidazole Elution buffer	5 M Tris-HCl 5 M NaCl 2 M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 150 ml Up to 1 L

Reagent/Buffer Solution	Composition	Quantity per litre dH_2O unless otherwise stated*
400 mM Imidazole Elution buffer	5 M Tris-HCl 5 M NaCl 2 M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 200 ml Up to 1 L
500 mM Imidazole Elution buffer	5 M Tris-HCl 5 M NaCl 2 M Imidazole Adjust pH to 7.4 ddH_2O	10 ml 100 ml 250 ml Up to 1 L
Ion Exchange Start buffer	5 M Tris-HCl Adjust pH to 7.4 ddH_2O	10 ml Up to 1 L
Ion Exchange 100 mM Elution buffer	5 M Tris-HCl 5 M NaCl Adjust pH to 7.4 ddH_2O	10 ml 20 ml Up to 1 L
Ion Exchange 200 mM Elution buffer	5 M Tris-HCl 5 M NaCl Adjust pH to 7.4 ddH_2O	10 ml 40 ml Up to 1 L
Ion Exchange 300 mM Elution buffer	5 M Tris-HCl 5 M NaCl Adjust pH to 7.4 ddH_2O	10 ml 60 ml Up to 1 L
Ion Exchange 400 mM Elution buffer	5 M Tris-HCl 5 M NaCl Adjust pH to 7.4 ddH_2O	10 ml 80 ml Up to 1 L
Ion Exchange 500 mM Elution buffer	5 M Tris-HCl 5 M NaCl Adjust pH to 7.4 ddH_2O	10 ml 100 ml Up to 1 L

2.4 Growth Media

Table 2.4 describes the growth media recipe. All media were autoclaved as described in section **2.3**.

Table 2.4: Growth media

Media	Ingredients	Quantity per litre
2 x YT	Tryptone	16 g
	Yeast Extract	10 g
	NaCl	5 g
	dH_2O	Up to 1 L
Luria Bertani (LB) Broth and Agar	Tryptone	10 g
	Yeast Extract	5 g
	NaCl	5 g
	dH_2O	Up to 1 L
	Agar (for solid media only)	10 g

2.5 Antibiotic selection

All antibiotics were made up as stock solutions at concentrations shown in table 2.5 using an appropriate solvent and filter sterilised if required. Solutions were stored at -20 °C; with the exception of Hygromycin which was stored at 4° C.

Working concentrations are also given as these differ between bacterial species.

Table 2.5: Antibiotics

Antibiotic	Stock conc. (mg/ml^{-1})	*E. coli* working conc. ($\mu g/ml^{-1}$)	*S. coelicolor* working conc. ($\mu g/ml^{-1}$)
Ampicillin	100	50	-
Apramycin	100	100	25
Thiostrepton	100	25	25
Kanamycin	25	25	25

2.6 Culture conditions

2.6.1 Growth and storage of *E. coli* strains

E. coli cells were grown using LB and 2xYT media at 37 °C in a static temperature controlled incubator if grown on solid media or shaken at 220 rpm if liquid media was used with use of relevant antibiotics for selection (if applicable). Glycerol stocks of *E. coli* strains were prepared from 5ml of overnight cultures. Cells were centrifuged and resuspended in 0.5 ml of sterile 20 % glycerol (v/v) and stored at -80 °C. Cells grown on solid or liquid media were placed at 4 °C for short-term storage.

2.6.2 Growth and storage of *S. coelicolor* strains

For growth of *Streptomyces coelicolor*, both liquid and solid media may be utilized, incubating the cultures at 30 °C and when applicable containing relevant antibiotics for selection. In the case of using a liquid media, for optimum aeration of the growth media vessels were shaken at 225 RPM. Spore or aerial mycelia suspensions of *S. coelicolor* strains were stored in microcentrifuge tubes at -20 °C

2.7 Transformation

2.7.1 Preparation of electro-competent *E. coli* cells

Overnight cultures of *E. coli* cells were used as an inoculum, diluting 1/100 into 50 ml of fresh LB media. Cultures were grown to a density where the OD_{600} reaches 0.5 before cells could be harvested by centrifugation at 4000 rpm for 10 min at 4 °C. The supernatant was then decanted and cells gently resuspended in 50 ml of ice cold sterile 10% glycerol. This step is repeated using 20 ml ice cold sterile 10 % glycerol and finally 10 ml ice cold sterile 10% glycerol before being resuspended in 1ml ice cold sterile 10 % glycerol and 40 µl aliquots of the competent cells were prepared in microcentrifuge tubes. These aliquots were stored at -80 °C.

2.7.2 Transformation of electro-competent *E. coli* cells

A 40 µl aliquot of electro-competent *E. coli* cells were thawed on ice prior to the addition of 2 µl DNA and gentle mixed by pipetting. After a brief incubation on ice, the cells are then transferred to a pre-chilled electroporation cuvette and subjected to electroporation using a BioRad MicroPulser. Immediately after electroporation,

1 ml of LB medium was added to the cuvette to collect cells and then transferred to a sterile universal tube to be incubated at 37 °C whilst shaking at 225 rpm. After 90 min incubation, the

culture was plated on LB agar plates with appropriate antibiotic selection and 1 % glucose where protein expression requires tight regulation.

2.8 DNA isolation and manipulation

2.8.1 Plasmid DNA isolation from *E. coli*

Small-scale purification of plasmid DNA was carried out using the AccuPrep® Plasmid Mini Extraction Kit (Bioneer) based on the alkaline lysis method described by Sambrook *et al.*, (1989). *E. coli* cells (usually JM109) from an overnight culture grown in 5 ml LB broth containing the relevant antibiotic selection for the plasmid were pelleted at 13000 rpm for 5 min and the supernatant removed. The cells were resuspended in 250 µl resuspension buffer prior to the addition of 250 µl lysis buffer and mixed by inverting the microcentrifuge tube 3-4 times. This was followed by the addition of 350 µl neutralisation buffer and mixing by inversion 3-4 times. The resulting suspension was centrifuged at 13000 rpm for 10 min and the supernatant was transferred to spin column containing a silica DNA binding matrix. DNA was bound to the matrix by centrifugation at 13000 rpm for 1 min and the flow through discarded. The matrix was washed with 750 µl wash buffer containing ethanol and the flow through discarded. The column was then dried with another centrifugation at 13000 rpm for 1min to remove excess ethanol that would interfere with down-stream applications. Finally, the DNA was eluted from the matrix into a sterile microcentrifuge tube using 50-100 µl sterile ddH_2O by centrifugation at 13000 rpm for 1 min.

2.8.2 Enzymatic manipulation of DNA

Restriction digestion of DNA was performed with enzymes from New England Biolabs (NEB) or Promega Corp and carried out in accordance to the manufacturer's instructions. Ligation reactions were set up by mixing vector and insert DNA at a ratio of 1:3 and incubated overnight at 16 °C with T4 DNA ligase and supplied buffer. DNA was diluted 1:1 with ddH_2O prior to transformation into electro-competent *E. coli* cells.

2.8.3 Preparation of Lambda HindIII DNA size ladder

Lambda DNA (NEB) was completely digested with *Hin*dIII restriction enzyme (NEB) to yield a reproducible molecular weight standard containing bands ranging from 125 bp to 23 kb in size. To a total of 50 µg of lambda DNA (NEB) in a 450 µl reaction buffered with NEB buffer 2, *Hin*dIII restriction enzyme (NEB) corresponding to 100 units (5 µl) was added and the reaction incubated at 37 °C for 90 min prior to the addition of an extra 5 µl *Hin*dIII enzyme.

Further 90 min incubation at 37 °C was followed by heat inactivation of the enzyme at 75 °C for 10 min prior to the addition of 50 µl of 6X DNA loading dye.

2.9 DNA quantification

2.9.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA according to size with the percentage of agarose used varied dependent on the size of DNA molecules of interest. 1% (w/v ratio) agarose dissolved in 0.8X TBE buffer containing ethidium bromide (final concentration of 0.1 µg/ml) was predominantly used to resolve DNA fragments in the region of 0.5 to 15 Kb. Gels were immersed in 1X TBE in BIORAD electrophoresis tanks and samples mixed with the 6X DNA loading buffer at a ratio of 5:1 sample to buffer. Samples were loaded into the wells of the gel and electrophoresis carried out at 100 V for a time relevant to the size of the DNA fragments of interest. Lambda *Hin*dIII molecular size ladder was often run in a spare lane for estimation of DNA size. Once run, gels were visualised and imaged utilising a BIORAD gel documentation system (GelDoc).

2.9.2 Purity and concentration of nucleic acids

Purity and concentration of nucleic acids were assessed using a NanoDrop ® ND-1000 spectrophotometer with a pathlength of 0.2 mm. Prior to measuring samples, the instrument was initialised and blanked using ddH_2O using a volume of 2 µl. Nucleic acid concentrations are measured at a wavelength of 260 nm, while purity is assessed by measuring the absorbances at 280 nm and 230 nm and creating 260:280 and 260:230 ratios. Ratios below ~1.8 may indicate contamination of the nucleic acid samples with proteins, ethanols or other compounds that may have been co-purified.

2.9.3 Site directed mutagenesis

Site directed mutagenesis was carried out using a QuickChange® Lightning Site-Directed Mutagenesis Kit from Agilent Technologies according to the manufacturers guidelines. Primers were manually designed to create a single point mutation flanked by unmodified nucleotide sequence. The resulting SDM creates a restriction enzyme recognition sequence within the mutated DNA sequence.

PCR reaction mixtures were set up in thin-walled PCR microcentrifuge tubes to a final reaction volume of 51 µl. The reaction set up was a follows

5 µl of 10x reaction buffer
100 ng of dsDNA template (in this case a plasmid)
125 ng oligonucleotide primer #1
125 ng oligonucleotide primer #2
1 µl of mixed dNTPs
1.5 µl QuickSolution reagent
Correct final volume up to 50 µl
Then add:
1 µl QuickChange® Lightning Enzyme

The PCR cycling parameters for the SDM method are given in table 2.6

Table 2.6: Site-Directed Mutagenesis cycling parameters

Segment	Cycles	Temperature	Time
1	1	95 °C	2 min
2	18	95 °C	20 sec
		60 °C	10 sec
		68 °C	30 sec/kb of plasmid (219 sec used)
3	1	68 °C	5 min

An extension time of 3.65 min was used.

Following PCR, 2 µl of the kits provided *Dpn* I enzyme was added directly to the reaction mix and incubated at 37 °C for 5 min to digest the template (non-mutated) supercoiled dsDNA.
The resulting mutated plasmid was transformed into *E. coli* JM109 for downstream cloning and confirmation.

2.10 Protein production

2.10.1 Overexpression in *Streptomyces coelicolor*

Over expression of *Sc*Dps was carried out in liquid cultures. Cultures were started with the inoculation of the media with a small volume (5 µl) of stock spore suspensions of relevant strains. The Dps plasmids were selected for with the addition of apramycin. 24-hour liquid cultures were spiked with thiostrepton (25 µg ml^{-1}) to initiate over expression and were incubated for a further hour. Cells were collected by centrifugation at 8000 rpm for 5 min and the waste media removed. Preparations of cell free extracts were created by cellular disruption via sonication followed by centrifugation to remove cellular debris. Buffers allowing native conditions to be preserved were used (20 mM Tris-HCl, 200 mM NaCl, pH 7.5).

2.10.1 Recombinant protein expression

A single colony from of plate of BL21 (DE3) cells containing the relevant plasmid was picked and grown overnight in 2 x YT / 1 % glucose and kanamycin for pET plasmid selection (37 °C, 250 rpm). The over-night culture was used to inoculate fresh 2 x YT media containing kanamycin but lacking glucose. The culture was grown until the OD reached 0.9 at 600 nm. Protein expression was then induced with the addition of 1mM IPTG and the cultures were then transferred to the 30 °C shaking incubator for 2 hr. The cells containing the protein of interest were collected by centrifugation at 8000 rpm for 5 min and the waste media removed. Cells were processed immediatey for protein purification or stored at -80 °C

2.10.2 Purification of His-tagged proteins

Immobilised metal ion affinity chromatography (IMAC) was used as a means of purifying His-tagged proteins from complex cell lysate mixtures. HisTrapHP™ (GE Healthcare) were utilised as these Nickel-Sepharose™ pre-packed columns provide a high performance means for fast and convenient preparative purification of His-tagged proteins. Two column sizes were used; 1 ml and 5 ml with the dynamic binding capacity at least 40 mg of histidine tagged protein per ml of column medium. Thus for the larger protein preparations the 5 ml column was commonly used.

A BL21 (DE3) cells pellet containing the expressed protein of interest was resuspended in binding buffer (50 mM Tris-HCl, 500 mM NaCl, 50 mM Imidazole pH 7.4) to achieve a concentration factor of X50 (20 ml per 1 L culture). Lysozyme was added to a concentration of 0.2 mg/ml to aid cell lysis and the cell suspension incubated at room temperature for 30 min and then on ice for 30 min. Sonication was then used to fully disrupt cells using a Branson Digital Sonifier at 70% amplitude for 1 min bursts keeping samples on ice at all times. The total protein extract was clarified, removing cell debris by centrifugation and the resulting supernatant passed

through a 0.2 μm syringe filter unit. The supernatant was applied to a HisTrapHP™ column pre-equilibrated with at least 5 column volumes of binding buffer. Once sample is fully applied, the column is washed using 10 column volumes of washing buffer. Flow through and column wash solutions are collected to confirm no major loss of protein has occurred during these steps.

A stepwise gradient of imidazole is used to elute the protein from the column using concentrations at 100 mM, 200 mM, 300 mM, 400 mM and 500 mM and at least 5 column volumes at each concentration collecting fractions of appropriate volumes. Samples of each fractions including the flow through and washing steps were assessed for purity and homogeneity by SDS PAGE followed by coomassie staining. Pure samples were pooled and buffer exchanged.

2.10.3 Protein buffer exchange

Size exclusion chromatography was used to change the buffer in which proteins were kept. HiTrap™ Desalting columns (GE Healthcare) were used for buffer exchange after purification, protein refolding and other steps such as enzymatic reactions. The 5 ml volume pre-packed column was equilibrated with 25 ml of required buffer at 5 ml/min prior to the application of a 1.5 ml sample volume and the 1.5 ml eluted buffer discarded. 1.5 ml of fresh buffer was then added and the 1.5 ml eluted from the column collected, this fraction contained the protein of interest. The column was rinsed with 10 ml of buffer before application of another 1.5 ml sample. These columns were scaled up for larger samples and volumes increased proportionally to the number of columns used (up to 5 inline columns maximum)

2.10.4 Ion exchange chromatography

Ion exchange chromatography was used for preparation of samples for crystallography studies. RESOURCE Q™ (GE Healthcare) 1 ml pre-packed columns were used. Samples were first buffer exchanged into the ion exchange start buffer (20 mM Tris-HCl pH 7.4) using the column desalting method previously described before application to the RESOURCE Q™ column pre equilibrated with start buffer at a flow rate of 4 ml/min. After application, the column is washed using 10 column volumes of start buffer before proteins can be eluted using a stepwise gradient of buffer with increasing ionic strength (100 mM, 200 mM, 300 mM, 400 mM and 500 mM NaCl). Fractions of 0.5 ml are collected and screened for homogenous protein via coomassie-stained SDS PAGE.

2.10.5 Protein quantification

A Bio-Rad protein assay kit based on the Bradford method was used to accurately determine the concentration of solubilised proteins. The dye reagent was freshly prepared by diluting the dye reagent concentrate with 4 parts ddH_2O to 1 part dye. BSA was used as a standard and prepared as shown in the following table

Table 2.7: Bradford test protein quantification standards

BSA (1 mg/ml)	ddH_2O	Final protein quantity
0 µl	20 µl	0 µg
2 µl	18 µl	2 µg
4 µl	16 µl	4 µg
8 µl	12 µl	8 µg
16 µl	4 µl	16 µg
20 µl	0 µl	20 µg

To the BSA protein standards, 1 ml of diluted dye reagent was added in a semi micro cuvette, mixed well by pipetting and incubated for 5 min. Optical density was recorded at 595 nm using the 0 µg solution as a blank. A standard curve was created using the OD values read from the BSA standard samples. Unknown protein solutions were measured in the same way but mixing 10 µl of sample with 10 µl ddH_2O and

1 ml diluted dye reagent and the concentration calculated using the standard curve straight-line equation.

The concentration of pure protein solutions was also quantified using a NanoDrop® ND-1000 spectrophotometer reading the absorbance at 280 nm using a reference setting based in a 0.1% (1 mg/ml) protein solution giving an absorbance of 1.0 at 280 nm. The spectrophotometer was initialised using ddH_2O and blanked using the protein's buffer. 2 µl sample volumes were used and the units given in mg/ml.

2.11 Protein separation

2.11.1 SDS Polyacrylamide gel electrophoresis

SDS PAGE is a technique used to separate and visualise denatured proteins according to their electrophoretic mobility. The gels were hand cast using the formulae given in table 2.3. Protein solutions were mixed with loading buffer and heated at 95 °C for 6 min to fully denature the proteins prior to loading into the wells of the gel. Electrophoresis was carried out using SDS PAGE running buffer at a constant voltage of 150 V for 1-2 hr depending on the size of the

protein of interest. A molecular weight size marker was also run (Precision plus Kaleidoscope. Bio-Rad)

2.11.2 Native PAGE

Native PAGE is a technique similar to the denaturing SDS PAGE however, no denaturing agents are involved thus preserving the native folding state of the proteins and protein complexes. Coomassie G-250 is included in the cathode buffer as a charge shift molecule, imparting a negative charge on the proteins without denaturing them, thus allowing proteins to be resolved according to their molecular weight. Gels were hand cast using the recipe given in table 2.3 and run using the native PAGE running buffer (anode and cathode buffers). Samples were loaded using the Native PAGE sample buffer (Invitrogen) and run at a constant 90 V at 4 °C for 1-2 hr.

2.11.3 PAGE gel staining

Gels were stained with Coomassie R250 staining solution for 30 min and subsequently de-stained using the destaining solution. The R-250 stain was recovered for future use prior to the addition of the destaining solution which was replaced with fresh destaining solution every 30 min until bands were intense and the gel back round was clear. These stained gels were then placed on a white board and imaged using a Gel Doc system (Bio-Rad).

2.11.4 Immuno-detection of His-tagged proteins (Western Blot)

This is process allows specific proteins to be detected utilising the specificity of antibodies against the protein or its translationally fused tag. Protein loading was normalised by loading equal quantities of total protein. Proteins resolved via native polyacrylamide gel electrophoresis were transferred from the gel to a PVDF membrane via a semi dry blotting process. The PVDF membrane (Amersham) was cut to the size of the resolving section of the PAGE gel and was activated by a brief rinsing in methanol followed by rinsing in ddH_2O. The membrane was then soaked in pre chilled transfer buffer for 10 min or until required. Thick blotting paper was also cut and prepared in parallel as with the membrane, also soaked in pre chilled transfer buffer for 10 min or until required. Once the native PAGE gel has finished running, i.e the proteins of interest have been separated, the gel was soaked in 1X SDS PAGE running buffer for 30 min prior to equilibrating in chilled transfer buffer for 10 min. A semi dry transfer system (Bio-Rad) was set up and the transfer cassette assembled on the anode plate. First a piece of blotting paper is laid down and excess buffer and air bubbles removed by rolling over with a glass rod. Next the membrane was placed on the blotting paper, followed by the gel and finished with another piece of blotting paper. Air bubbles and excess buffer were removed at each stage with the glass rod. The transfer apparatus was assembled and transfer run at 20 V

for 25 min. After transfer the membrane was placed in the 10% blocking solution and gently agitated at room temperature for 1 hr. Following this first blocking step, the membrane was washed in TBS/Tween for 5 min and repeated three times. Anti-His HRP (Horse radish peroxidise) conjugate antibody (Qiagen) was diluted 1:5000 in the 5% blocking solution and following the washing steps, the 5% blocking/antibody solution was added to the membrane and incubated for 1 hour at room temperature. Four final washes (5 min each wash) were carried out using TBS/Tween to remove excess antibody and the western blot was then visualised utilising a chemiluminescence kit (ECL Advanced system from Amersham biosciences) and a Chemi-doc system (Bio-Rad).

2.11.5 Two-dimensional PAGE

Two-dimensional PAGE was carried out, but not in the form of standard of isoelectric focusing followed by SDS PAGE. Instead the 2D approach utilised a 1st dimension of non-denaturing native PAGE followed by a second dimension of denaturing SDS PAGE. Native and SDS PAGE gels were run as previously described, however the native PAGE gel was run containing replicated lanes. After running, the gel was cut in half to leave one half to stain and the other to process for the second dimension. Whilst destaining one half of the gel, the other half was incubated in SDS PAGE running buffer to start protein denaturation. The second dimension was run in two ways. Firstly a whole replicated lane was cut vertically and flipped 90° on top of a hand cast SDS PAGE resolving gel. Filter paper containing a size marker was also introduced to one end of the strip. SDS PAGE was carried out as previously described and stained with coomassie. Secondly the stained 1D gel was aligned with the unstained gel and the region where the band sits in the stained gel was cut from the unstained gel. This gel fragment was then cut into small strips and placed into a well of a hand cast SDS PAGE gel and run and stained as described.

2.12 Analysis of Dps mediated iron oxidation and deposition

2.12.1 Spectrophotometric analysis of Dps mediated protein oxidation

The progress of Dps mediated ferrous iron oxidation was followed at 25 °C monitoring the formation of the ferric hydrous oxide micelle by following its absorbance at a wavelength of 310 nm. Where oxygen was used as an oxidant, aerated reaction buffers were used, while hydrogen peroxide required a degassed reaction and was added at a half concentration of iron. Reactions were carried out in either a stepwise manner with sequential additions of ferrous iron and hydrogen peroxide when required as an oxidant, or single stage addition of substrates to their final reaction concentrations.

The reactions were monitored at 10 s intervals over the time course of the experiment. Fe^{2+} auto oxidation was monitored in parallel with reactions lacking protein.

2.12.2 Staining for Dps protein bound iron

Iron loaded Dps resolved by native PAGE were stained for ferric iron using the protocol of (Leong et al., 1992). Following electrophoresis, gels were incubated in potassium ferricyanide solution (100 mM potassium ferricyanide, 50 mM Tris-HCl, 100 mM NaCl, pH 7.5) in the dark for 10 min prior to destaining with a 10% methanol, 10% tri-chloroacitic acid solution prepared immediately before use. This destaining solution was changed every hour until required colour development, after which gels were imaged and then stained with coomassie R250.

2.12.3 DNA oxidative protection

In vitro DNA protection against oxidative damage was assessed using 40 ng pUC18 (2686 bp) in 20 mM Tris, 200 mM NaCl, 5% glycerol in a final volume of 20 µl. Dps protein was added to a final concentration of 0.5 µM and the DNA/protein mixture incubated at room temperature for 10 min prior to the addition of $[NH_4]_2[Fe][SO_4]_2 \cdot 6H_2O$ to reach 50 µM, subsequently, incubation for 5 min was followed by addition of H_2O_2 in molar excess (5 mM) and a further incubation for 10 min before reactions were resolved with a 1% agarose gel and stained with ethidium bromide. DNA protection assay was also carried out for protection against copper mediated damage using a final concentration of 50 µM $CuSO_4$ followed by 1 mM H_2O_2 5 mins later. Reactions were resolved as described previously after 10 min incubation time.

2.13 Protein structure resolution

Proteins were sent out of house for their structure to be resolved by X-Ray crystallography (sections 2.13.1 to 2.13.3). The PDB files were returned and validation was carried out before preforming analysis of the structures on site (sections 2.13.4).

2.13.1 Crystallisation

Sitting drop vapour diffusion 96 well plate screens were made using an Innovadyne Screenmaker 96+8, and the JCSG-plus HT96 screen (Molecular Dimensions). Sitting drops included 100 nL of precipitant with 100 nL protein, and 100 nL of precipitant with 200 nL protein. Best crystals formed at room temperature within 2 weeks in the following conditions C10 (0.1 M Bicine pH 9.0 with 10 % w/v PEG 20,000/ 2% v/v Dioxane) for DpsADCT, E2 (0.2 M sodium chloride, 0.1 M sodium cacodylate, 2.0 M ammonium sulfate pH 6.5) for DpsA and G12 (0.1 M Bis Tris pH 5.5 with 3.0 M sodium chloride) for DpsC.

2.13.2 X-Ray Data Collection and Structural Refinement

Crystals were cryoprotected with 25% v/v glycerol and frozen in liquid nitrogen. Diffraction data was collected on beamline I02 (DpsC) and I04 (DpsA and DpsADCT) at Diamond Light Source (Harwell Science and Innovation Campus, Didcot, Oxford, UK). Datasets were integrated and scaled with XDS (Kabsch, 2010) and XSCALE (Kabsch, 2010), through xia2 (Winter, 2010); the in house automatic data processing software at Diamond light source.

2.13.3 Structure determination

A search model was built using Clustal omega (Sievers et al., 2011) to align and Chainsaw (within ccp4 package) (Stein, 2008) to chain-trace *Streptomyces coelicor* DpsA (Q9R408) onto the *Mycobacterium smegmatis* Dps structure (1UVH). The structure of DpsA was solved by molecular replacement with Phaser (McCoy et al., 2007) using the aforementioned search model. The structures of DpsADCT and DpsC were solved using the refined structure of DpsA as a search model. The structures were refined iteratively with Refmac5 (Skubák et al., 2004) and model building in Coot (Emsley et al., 2010).

2.13.4 Online Structure Validations

PDB files were submitted to the ERRAT server found on the Protein Structure and Validation Server (http://services.mbi.ucla.edu/SAVES/) (Colovos and Yeates, 1993) for the identification of mistraced protein regions due to errors in model building.

Ramachandran plot analysis were carried out using RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php; (Lovell et al., 2003) submitting the full PDB file and downloading the output.

2.13.5 Proteins, Interface, Structure and Assemblies Analysis (PISA)

Full atomic co-ordinates of DpsA and DpsC in PDB format were uploaded, independently, to the PDBePISA interactive tool located at http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html (Krissinel and Henrick, 2007). The cell parameters and space symmetry group boxes were cross referenced, process ligands boxes unchecked, processing mode was set to automatic and interfaces analysis selected. The results included detailed statistics of the interface summery and residues involved in bonding downloaded and used to reconstruct images.

2.13.6 Visualisation, analysis and biocomputational generation of protein molecular structures and surfaces

UCSF Chimera (Pettersen et al., 2004) was used as a platform to visualise and image the molecular structures of the Dps proteins from their PDB files. All molecular images were created with this software using the embedded commands found within the program for adding surfaces, creating distance calculations, searching for H-bonds, superpositioning of proteins, generating electrostatic surface potentials and hydrophobicity maps.

2.13.7 Computational Alanine Scanning of Protein-Protein Interfaces

A computational approach was used for the prediction of energetically important amino acid residues involved in protein interfaces. The server (available at http://www.robetta.org/alascansubmit.jsp) requires a three dimensional structure in the form of a PDB file. Whilst the pipeline has a facility to predict residues involved in the interface, the polypeptide chains that form the interface are required to be defined. Thus in the case of DpsA and DpsC, individual submissions were utilised for the dimer interfaces, "ferritin-like" interfaces and "Dps-like" interfaces. The software predicts which residues involved in the interface would destabilise the interface if mutated to alanine. The main output result that confirms if a residue is a "hotspot" is $\Delta\Delta G$ complex score. All residues within the interface are assigned a $\Delta\Delta G$ complex score, but only values greater than 1 kcal/mol are considered destabilising and thus these residues are "hotspots".

Chapter 3

Recombinant Expression and Biochemical characterisation of *S. coelicolor* Dps proteins

3.1 Exploration of *S. coelicolor* Dps protein assembly *in vivo*

This chapter explores the basic biochemical features of the *Sc*Dps proteins; from their oligomeric assembly through to their ferroxidase activity and their functional significance in the ability to protect DNA. Previous research into the *in vivo* expression profiles of the *Sc*Dps proteins indicated that they have roles in stress response and reproductive growth (Facey et al., 2009). The ability of DpsA to assemble into a dodecamer *in vivo* was also investigated although little attention was paid to the properties of DpsB and DpsC. Thus this first chapter starts by investigating these properties *in vivo*.

As multimeric proteins, monitoring the molecular weight as a function of the protein's hydrodynamic radius can assess the oligomeric assembly of Dps proteins. In order to do this, Blue-Native polyacrylamide gel electrophoresis (BN-PAGE) was employed. This is a technique that allows an electrophoretic separation of proteins according to their size in native, non-denaturing conditions thus preserving any tertiary and quaternary structures that give proteins their shape and size. G-250 coomassie dye forms non-covalent electrostatic interactions with proteins in solution, binding predominantly to the positively charged basic amino acids. Consequently the use of a coomassie G-250 additive in the sample buffer and in the cathode buffer allows all proteins to carry a negative charge and therefore separation is based only on the hydrodynamic radius of a protein and not on its native charged state. This technique combined with immuno-detection by means of western blot allows for detection of the proteins within a total cell extract resolved by native PAGE.

In order to establish the *in vivo* oligomeric assemblies of the three *Sc*Dps, an over-expression system was used to produce adequate quantities of protein. The thiostrepton inducible promoter (*tipA*) was used to control expression of the *dps* genes to which a C-terminal hexa-histidine translational fusion was also included. The plasmids pDpsA9, pDpsB9 and pDpsC9 plasmids were individually introduced into *Streptomyces coelicolor* (A3) 2 M145 (Facey et al., 2009) by intergenic conjugation. This created three strains from which each protein could be over-expressed independently. Following induction of expression as described in section **2.10.1**, cell free extracts were prepared and equal volumes were resolved by BN-PAGE. Following a slightly amended blotting procedure which includes an incubation in SDS PAGE running buffer (**2.11.4**), immuno-detection was carried out using an anti-His HRP conjugate antibody (Qiagen) against the His-tag of the proteins. Chemiluminescence was achieved using an ECL Advance Western blotting detection kit (GE Healthcare) and was detected with a BioRad gel documentation system. The results indicated that over expression using the *tipA* promoter does not affect the ability of DpsA to assemble into putative dodecamers and mirrors the results observed under stress conditions (Facey et al., 2009). There

is also a signal detected from DpsA indicative of a lower, non-assembling species. DpsC also displays an assembly of dodecameric nature but, unlike DpsA, no low oligomeric species were detected. In contrast to DpsA and DpsC, DpsB fails to assemble as a dodecamer under these conditions producing only a non-assembling protein species with a mass smaller than the mass of DpsA's non-dodecameric species (Figure 3.1).

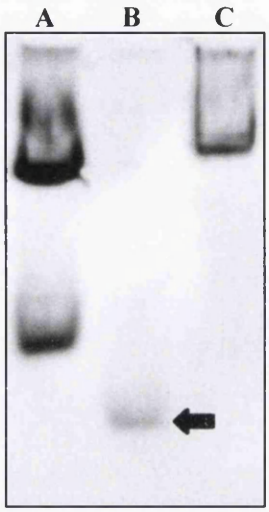


Figure 3.1: Immuoblot of BN-PAGE assessment of oligomeric assembly of DpsA, DpsB and DpsC from *in vivo* preparations. DpsA (lane A) and DpsC (lane C) both assemble however highlighted by the arrow is the non-assembling DpsB.

3.1.2 Construction of pET expression plasmids

In order to achieve *in vitro* characterisation of proteins, the coding DNA sequence (CDS) of the gene must first be cloned into a suitable expression vector and subsequently expressed using a host that can facilitate recombinant protein production using the expression system of choice. The genes encoding *Streptomyces coelicolor dpsA* (SCO0596), *dps*B (SCO5756) and *dpsC* (SCO1050) were extracted from pDpsA4, pDpsB4 and pDpsC1 respectively (Facey et al., 2009) by digesting with *Nde*I/*Bgl*II. Subsequent to resolving the digests by electrophoresis, the bands corresponding to the *dps* genes were extracted from the 1 % agarose using a commercial kit (GE Healthcare) and were then ligated into pET-26b(+) pre-cut with *Nde*I/*BamH*I. This creates three plasmids; pDpsA14, pDpsB14 and pDpsC14. Each contained translationally fused C-terminal 6X-His tag (Figures 3.2 and 3.3). The *dpsA* gene was also cloned into a *Nde*I/*BamH*I digested pET-16b vector for N-terminal poly-His tagging. Cloning was confirmed by restriction digest analysis and commercial Sanger sequencing using the T7 promoter and T7 terminator universal primers (LGC Genomics). The sequenced plasmids were transformed into *E. coli* K12 JM109, a strain that alleviates plasmid instability during cloning and preparation of plasmid stocks. This strain is not lysogenized to contain the T7 RNA polymerase required to cause expression of the target genes. Following sequence

confirmation, the plasmids were also transformed in to *E. coli* BL21 (DE3) for recombinant protein expression.

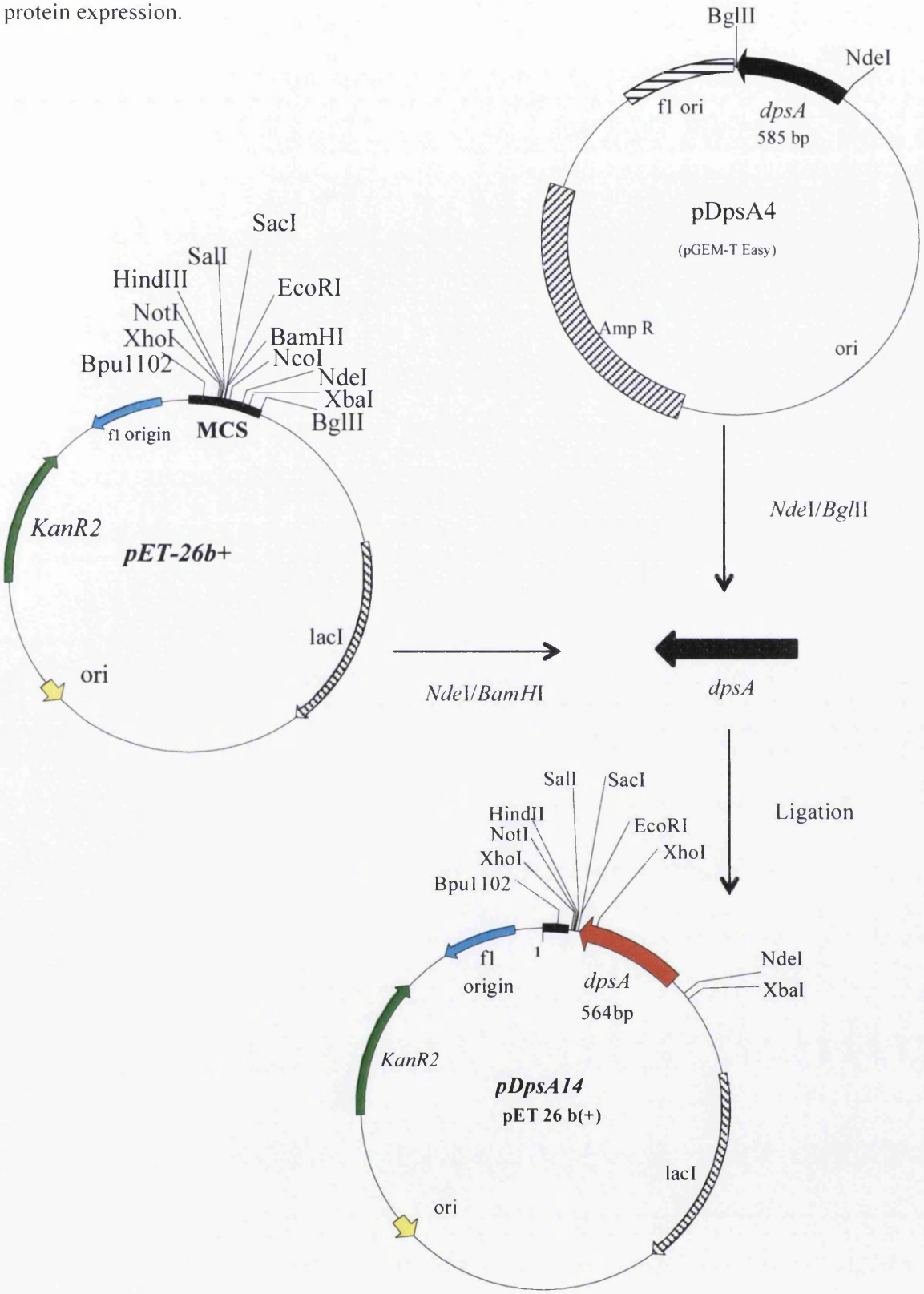


Figure 3.2: The cloning steps involved in creating the pDpsA14 plasmid. The *dpsA* CDS was excised from pDpsA4 digesting with *NdeI*/*BglII* and extracted from an agarose gel (1 %) before being ligated into a pET-26b(+) plasmid pre-cut with *NdeI*/*BamHI*. Sanger sequencing confirmed all sequences. Maps not drawn to scale.

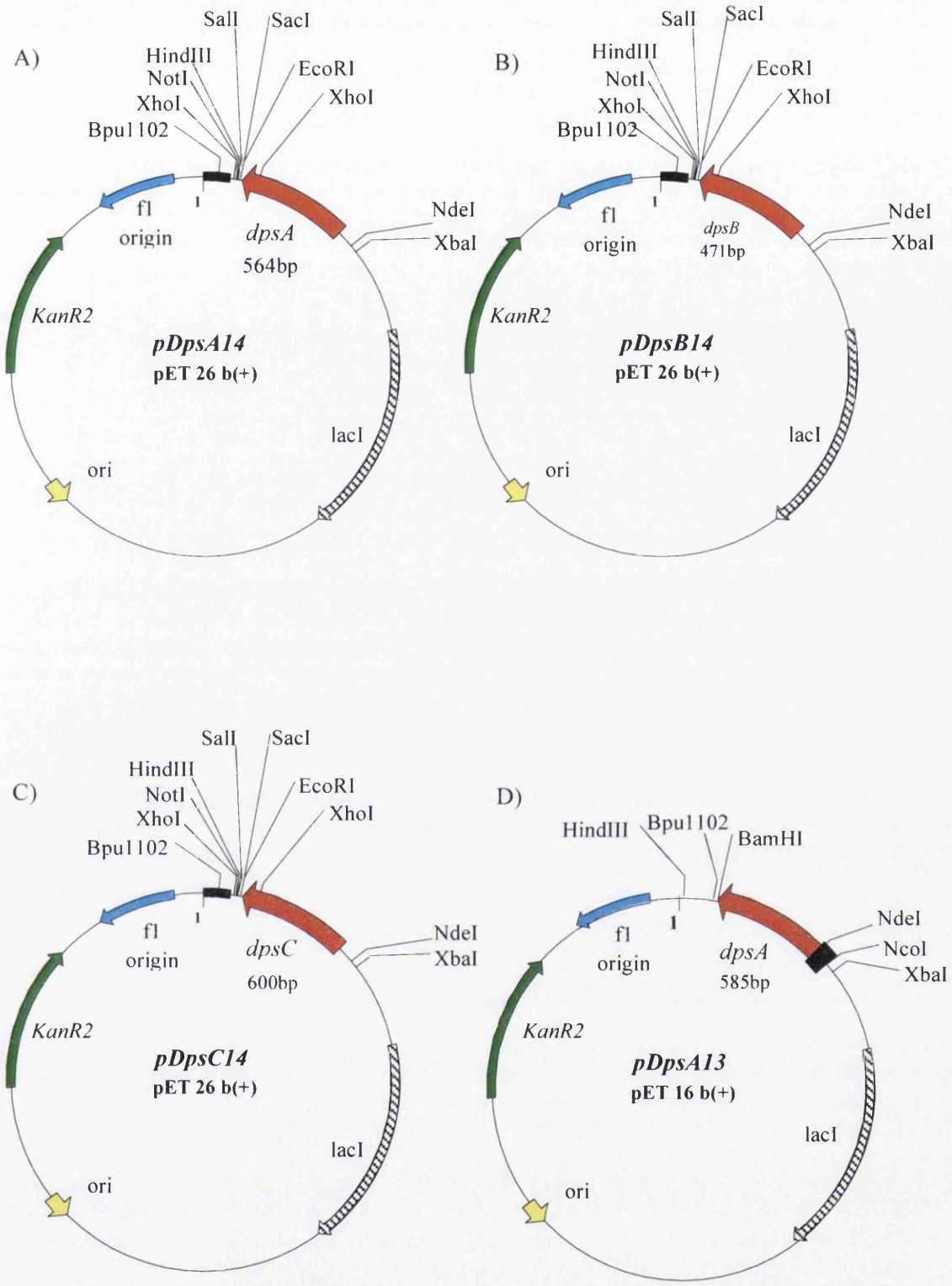


Figure 3.3: T7 based expression plasmid maps. *dpsA* (SCO0596) (A) and (D), *dpsB* (SCO5756) (B) and *dpsC* (SCO1050) (C) cloned into pET-26b (+). (A), (B) and (C) contain a C-terminal 6X-Histidine tag (D) contains an N-terminal 10X-Histidine tag. Maps not drawn to scale.

3.1.3 Recombinant protein expression and purification

Recombinant expression of the Dps proteins from *S. coelicolor* was carried out using the pET system of recombinant expression. This system utilises the bacteriophage T7*lac* promoter for transcription control; T7 RNA polymerase strongly selects and actively transcribes genes under the control of this promoter in order to produce high yields of protein upon induction of the system. An additional transcription control is located in the BL21 DE3 strain's chromosome; the *lac I* gene. This gene codes for the lac repressor protein which blocks "leaky" transcription from both the *lac* promoter, for T7 RNA polymerase transcription, and the T7 promoter used for target gene expression via a *lac* operator element. Glucose is required for the expression of *lac I* and thus is used in the media as a means of blocking target gene expression.

E. coli BL21 (DE3) can accumulate two forms of over-expressed recombinant protein; insoluble protein aggregates known as inclusion bodies or soluble proteins. The former is often composed of partially or mis-folded denatured proteins that lack biological activity, while the latter is typically accounted for predominantly by correctly folded and biologically functional protein. Since there were almost no documented examples of problems encountered with the solubility of recombinantly produced Dps proteins, the protocol used for expression of *S. coelicolor* Dps proteins adhered to those reported in the literature and is described in section **2.10.1**. The expression conditions used were to facilitate the production of soluble, active protein. Moreover this allows for purification in non-denaturing conditions utilising only the soluble fraction of cell extracts for downstream purification. Since Dps proteins are associated with an ability to bind DNA (Grant et al., 1998) and accounts of large DNA/Dps complexes have been reported (Ceci et al., 2004), the use of sonication was employed not only to lyse cells and release the intracellular content including the recombinant protein, but also to shear DNA and prevent potential formation of large DNA/Dps aggregates that could pellet out during the separation of insoluble matter from soluble matter via centrifugation.

The expression of soluble protein for each *dps* construct in BL21 (DE3) cells was successful enough to purify relatively high quantities of protein, yielding on average 150 mg of pure target protein per litre of culture. After clarification of the crude lysate via centrifugation, immobilised metal ion affinity chromatography (IMAC) was used to purify the proteins using their C-terminal histidine tags. In order to fully assess the purity of the eluted fractions, 15 µl sample volumes were assayed by SDS PAGE and the gels stained with coomassie R250 (Figure 3.4A). This results in very large target protein bands, and it also allows visualisation of bands from contaminating proteins that would otherwise remain unseen if smaller loading volumes were used. This purification method removed a great deal of contaminating proteins that were lost in the flow through of the column, not being able to bind to the nickel charged Sepharose in

the low imidazole concentration of the binding buffer. Washing using this buffer with over 10 column volumes also removed additional unwanted protein. However this first stage of purification via IMAC resulted in the elution of the target proteins over quite a large range of aliquots where elution started at 100 mM imidazole concentration and continued until all the bound protein was removed in the latter fractions of the 500 mM imidazole aliquots. This broad range of elution meant the fractions containing target protein also contained other contaminating proteins. Nevertheless pure proteins were found in aliquots at the higher concentrations of imidazole (400 – 500 mM), albeit at lower protein concentrations. This rapid initial capture stage of purification can be complemented with a further purification and cleaning step in order to achieve high concentrations of a homogenous protein solution. Ion exchange chromatography is commonly employed for this purpose and since DpsA has a predicted pI of 5.40, DpsB has a predicted pI of 4.63 and DpsC a pI of 5.90; it was considered that an anionic system would best suit the task. RESOURCE™ Q was used which requires buffers to be at least 0.5 to 1.0 pH units above the pI of the protein. An ionic strength of zero was used to apply the protein solution to the column, which required the purified protein solutions to be buffer exchanged to a buffer containing no salt (20 mM Tris-HCl, 5% glycerol pH 8.0). This was achieved using the HiTrap Desalting columns, a method based on gel filtration. Ion exchange chromatography was carried out as described in section **2.10.4**. The process resulted in the removal of virtually all-contaminating proteins from the samples (Figure 3.4B). Further buffer exchange was then required to return the proteins to an appropriate ionic strength buffer (20 mM Tris-HCl, 200 mM NaCl, 5% glycerol and pH 7.5). When high protein concentrations were required, Amicon® Ultra centrifugal filter units with a molecular weight cut off (MWCO) of 20 kDa were used to concentrate the protein solutions.

3.2.1 Tag position affects oligomeric assembly in recombinant DpsA

DpsA13 (N-terminal His-tag) and DpsA14 (C-terminal His-tag) were purified to homogeneity and subjected to BN-PAGE on a 7 % home cast gel. Horse spleen ferritin (Sigma), with a molecular weight of 440 kDa, was used as a rough size marker (de Val et al., 2012). DpsA13 with the N-terminal tag was the first construct to be tested, and whilst expression and non-denaturing purification yielded soluble protein, its size estimation by BN-PAGE showed that its assembly was not as expected. A smear like band was consistently seen from different protein preparations with a mass above that of ferritin (440 kDa) (Figure 3.4C). There could be two reasons for this result; either uncharacteristically high oligomeric assembly occurs with DpsA or the placement of the N-terminal histidine tag interferes with correct tertiary or quaternary protein structure resulting in protein aggregates. This type of assembly was not detected *in vivo*, however the pDpsA9 construct codes for DpsA protein translationally fused

with the His-tag at the C-terminal and thus points toward the N-terminal tag causing steric hindrance in the folding or multimeric assembly of the protein.

Indeed, purified DpsA14 tagged at the C-terminal resolved as a compact band of smaller size than ferritin. This indicates the correct assembly of the protein into a dodecamer. There is also the presence of lower oligomeric assemblies, which again mirror the data generated from the *in vivo* studies. What must also be taken into account when analysing the migration with in a PAGE gel is the actual molecular weight/protein external diameter ratio. The horse spleen ferritin with a molecular weight of 440 kDa has an external diameter of 12-13 nm (Harrison and Arosio, 1996). Dps proteins are said to have an external diameter of roughly 10 nm but have average molecular weights of around 260 kDa (Haikarainen and Papageorgiou, 2010). Thus, simply according to the molecular weight of DpsA it should migrate further in the gel, however its actual dimensions prevent this.

From this data it can be concluded that correct oligomeric assembly of DpsA is disrupted by the placement of a tag at the N-terminal region of the protein. The exact manner in which the tag interferes with assembly cannot be determined from these results, however the C-terminal histidine tag appears not to affect assembly and so points to the importance of the N-terminal tail in oligomeric assembly and suggests a role for proper dodecamer formation.

3.2.2 *In vitro* oligomeric assembly of DpsB and DpsC

Recombinant DpsB, expressed and purified was also subjected to BN-PAGE to assess oligomeric assembly. However, unlike DpsA, DpsB migrated to the lower region of the PAGE gel just as observed with the *in vivo* experiment. This again indicates that it cannot assemble into a dodecamer under these experimental conditions (Figure 3.4D). DpsC on the other hand does show a multimeric assembly and in accordance with it having a higher monomeric molecular weight than DpsA, the high oligomeric assembly is also slightly larger than DpsA suggesting that DpsC can too form a dodecamer (Figure 3.4D). The lack of assembly of DpsB dodecamers determined that its study was not continued any further.

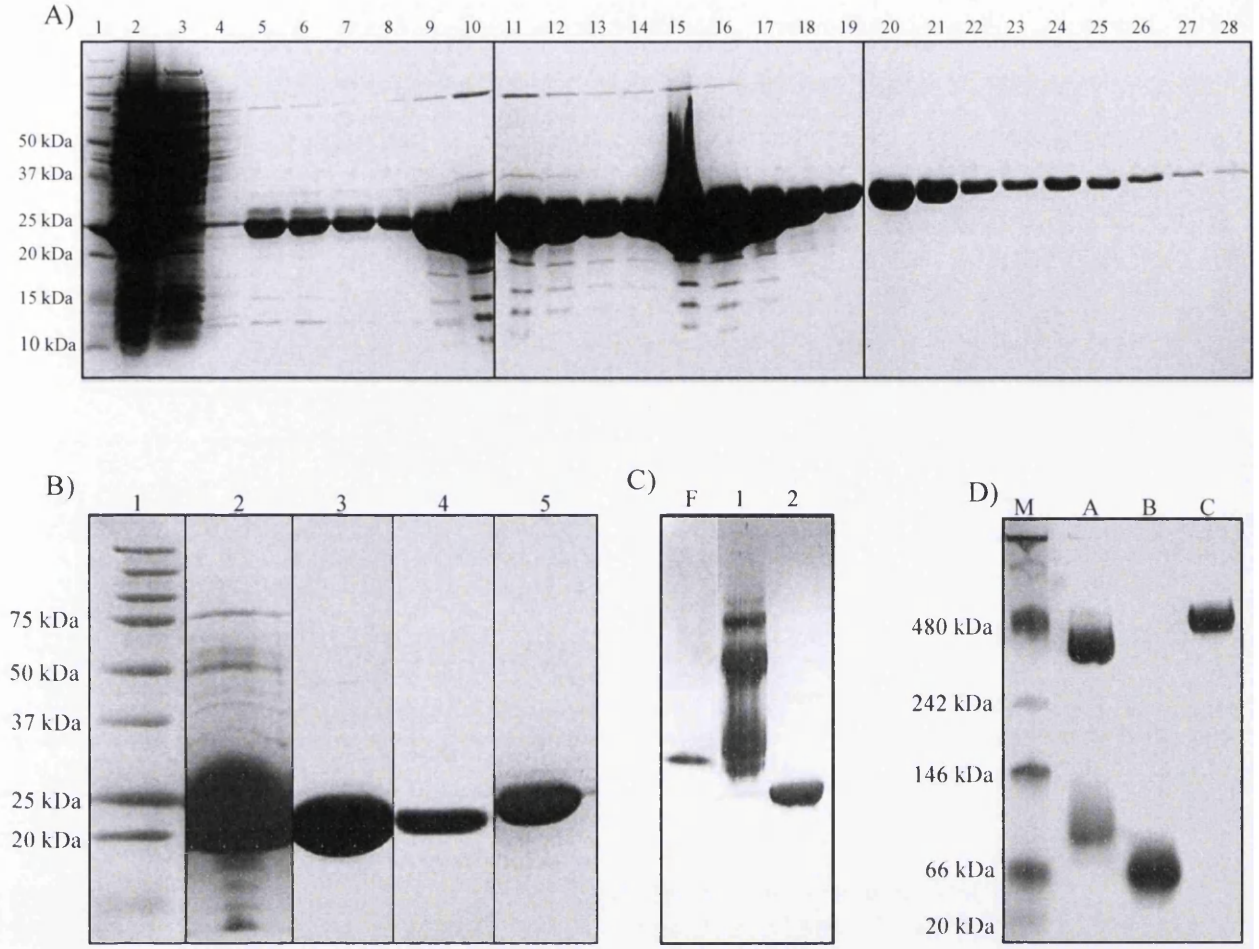


Figure 3.4: Purification and oligomeric assembly assessment of recombinant *Sc*Dps

(A) The C-terminally tagged DpsA was purified using IMAC with a stepped gradient of imidazole concentration; composite image of SDS PAGE gels (15 %) with fractions from the purification procedure run in each lane. Lane 1 – Biorad Precision Plus Protein™ Kaleidoscope™ standard, lane 2 – total soluble protein, lane 3 – total column flow through, lane 4 – total wash, lanes 5-9 (100 mM imidazole), lanes 10-14 (200 mM imidazole), lanes 15-19 (300 mM imidazole), lanes 20-24 (400 mM imidazole), lanes 25-28 (500 mM imidazole).

(B) Composite image of SDS PAGE gels used to assess samples after anionic exchange chromatography. Lane 1 – Biorad Precision Plus Protein™ Kaleidoscope™ standard, lane 2 – IMAC purified DpsA14, lane 3 – ion exchanged DpsA14, lane 4 – ion exchanged DpsB14, lane 5 – ion exchanged DpsC14.

(C) Composite image of native PAGE gel stained with coomassie R-250 highlighting how the placement of the poly His tag influences assembly of the DpsA dodecamer and indicates a role for the N-terminal tail in oligomeric assembly; lane F – horse spleen ferritin (440 kDa), lane 1 – DpsA13 (N-terminal His tag) and lane 2 – DpsA14 (C-terminal His tag).

(D) The *in vitro* oligomeric assembly of DpsA, DpsB and DpsC demonstrated by native PAGE stained with coomassie R-250 and running a NativeMark™ standard (Invitrogen). DpsA14 (lane A) and DpsC (lane C) both assemble into dodecamers while DpsB (lane B) does not.

3.3 Ferroxidase activity

Dps proteins possess the ability to oxidise ferrous iron into ferric iron to form a ferrihydrite core of nano scale. This process simultaneously utilises oxidants such as hydrogen peroxide (H_2O_2), thus sequesters both free iron and reactive oxygen species from the environment. The formation of the ferric iron core can be followed spectrophotometrically by monitoring its absorbance at 310 nm, a wavelength at which the ferrous iron does not absorb. Two oxidants are known to be utilised for oxidation in Dps proteins: hydrogen peroxide and oxygen, although the potency of these oxidants does differ with hydrogen peroxide regarded as the most effective for most but not all Dps (Liu et al., 2006). Because of the requirement of the oxidant this reaction can be considered a multi-substrate process since limitation of the oxidant impacts on the reaction rate. The inability of DpsB to form multimeric assemblies meant it was not included in these ferroxidase activity-monitoring experiments.

3.3.1 DpsA can utilise hydrogen peroxide and oxygen for ferroxidase activity

The iron oxidation progress was first monitored using DpsA and hydrogen peroxide (Figure 3.5A). Fully degassed buffers were used when performing oxidation with H_2O_2 since oxygen may also be utilised by the protein. In addition an auto oxidation process also occurs even in the absence of protein and can utilise excess oxidants attributing 20 – 25% of total oxidation in a complete system (Bakker and Boyer, 1986). To a 0.2 mg/ml solution of DpsA (746.27 nM), ammonium iron (II) sulphate was added to a final concentration of 100 μM and the reaction monitored for 1 minute prior to the addition of hydrogen peroxide (50 μM). Prior to the addition of H_2O_2, very little absorbance was recorded, however 10 seconds after the addition of hydrogen peroxide the reaction yielded an almost maximum absorbance suggesting the reaction had completed and thus its progress was missed (Appendix 1). A control reaction in the absence of protein was also conducted and exhibited a conversion of ferrous to ferric iron forms. This process is auto-oxidation and occurred at a rate slower than the reaction with protein. However, visual inspection of the reaction mixtures approximately 60 minutes post reaction highlighted how the ferric iron present in the control reaction precipitated out of solution and was present as sediment on the bottom of the cuvette. The protein solutions however, maintained the "rusty" colour of ferric iron indicating that most of the iron is stored within the protein cage and is maintained within solution. Nevertheless what cannot be ignored is the fact that there is likely be competition between protein mediated oxidation and auto oxidation, and each process may lead to a slowing of the other. Moreover certain reaction conditions have been shown to favour the auto oxidation process typically resulting from high Fe (II), high Fe (III) and high oxidant concentrations (Yang and Chasteen, 1999), showing how high substrate concentrations should be avoided.

DpsA was also tested with oxygen as an oxidant (Appendices 2-6), indicating that whilst not as potent as hydrogen peroxide and accordingly having slower ferric ion formation; oxygen was still capable of being utilised as an oxidant. Auto-oxidation was significantly lower than protein mediated oxidation and so can be considered to have little impact on the protein reactions, unlike when hydrogen peroxide was employed as an oxidant. Figure 3.5B shows the progress curves for iron oxidation using oxygen. The reactions were initiated upon the introduction of deaerated ferrous iron to the aerated reaction mixture (time point 0). Following the reactions the iron remained in solution, giving an early indication that it was deposited within the cavity of DpsA. However calculation of the concentration of ferric iron present within the protein solutions proved problematic. Unbound ferric iron is required to generate an extinction coefficient in order to calculate concentration of iron within solution. However the unbound ferric iron behaves differently in solution to the ferric iron deposited within the cavity of the Dps; free ferric iron is not soluble. Nearly all Dps papers within the literature contain iron oxidation curves that display the raw graphical absorbance data over time and do not convert these arbitrary units to absolute concentration values as the error involved is too high (Bhattacharyya and Grove, 2007), (Gupta and Chatterji, 2003, Alaleona et al., 2010).

3.3.2 DpsC ferroxidase activity; DpsC can utilise hydrogen peroxide

Since the oxidation rate of DpsA using hydrogen peroxide was too rapid to monitor, DpsC mediated oxidation was carried out utilising oxygen as an oxidant. Conditions were kept as described for DpsA, however upon addition of the ammonium iron (II) sulphate at a concentration of 138 μM, DpsC precipitated out of solution. Fresh preparations of both protein and iron solutions were tested, nevertheless problems with protein precipitation was still encountered. In order to alleviate this problem, different reaction conditions were analysed, using both high (500 mM), low (10 mM) concentrations of sodium chloride however this did not improve the solubility of DpsC when iron was added to the reaction.

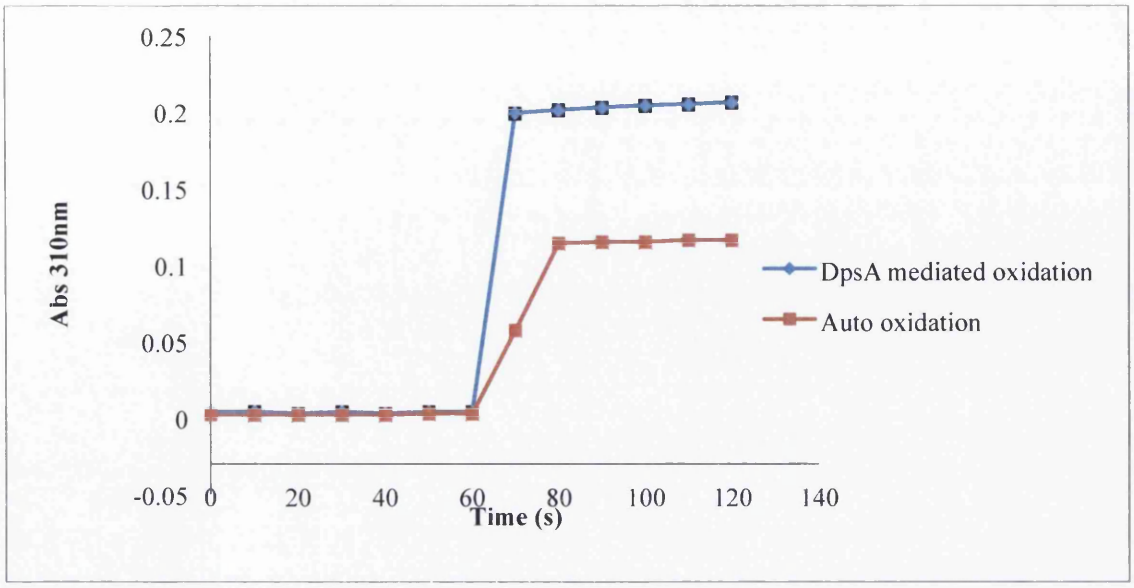
Hydrogen peroxide, as a more efficient oxidant, was tested. To avoid complete addition of ferrous iron into the reaction at a single stage, addition was carried out in a stepwise manner supplying 23 μM of iron per step. Hydrogen peroxide was added to the reaction mixture 1 min prior to the addition of iron (II) so that it was instantly available at the ferroxidase centre within the cavity of the protein. Additions of H_2O_2, each to a final concentration of 12 μM, were independently followed by an addition of iron (II) so that a total of 6 additions of each were made. This removed the problem with protein precipitation. Figure 3.6A highlights how addition of H_2O_2 produces no increase in absorbance, yet addition of ferrous iron produces a rapid increase in absorbance at 310 nm, analogous to what was seen with DpsA (Appendix 7). One minute prior to the next addition of Fe (II), further H_2O_2 was added. A lack of absorbance

increase indicated that the iron from the first addition was all successfully oxidised and that the oxidant was not a limiting factor. The pattern of iron addition and increase in absorbance of approximately 0.06 continued for the entirety of the reaction indicating that neither ferrous iron nor hydrogen peroxide were limiting during the reaction steps. This demonstrates how DpsC mediated oxidation is dependent on H_2O_2 and that the use of O_2 not only retards oxidation but the inability to oxidise the iron causes the protein to precipitate. This is a result that mirrors that of the Dps from *S. solfataricus* (Wiedenheft et al., 2005) whereby only H_2O_2 is employed to efficiently oxidise ferrous iron.

3.3.3 Ferric iron mineralisation

In order to confirm if the Dps proteins were capable of mineralising and depositing the ferric iron within their cavities as an iron oxide core, *in vitro* in gel ferric iron staining was used. This technique provides a method for detecting nonheme iron proteins which have been resolved in polyacrylamide gels via electrophoresis in native, non-denaturing conditions and relies on an acidic environment for the reaction of the protein bound iron with potassium ferricyanide (Leong et al., 1992). The stain is described as being sensitive enough to detect ferritin down to 1 μg of protein within the gel. However, there is a vast disproportion in quantity of iron capable of being held within the cavities of ferritin and Dps with almost a 10 fold decrease in iron within Dps when compared with ferritin; circa 4500 atoms are capable of being held within the ferritin cavity when compared with a maximum of 500 atoms within Dps (Crichton and Declercq, 2010). Because of the difference in iron content, 20 μg of iron loaded Dps was resolved on a 7.5% native PAGE gel. Due to the stain resulting in a blue band, clear native PAGE was performed avoiding the addition of the coomassie G-250 additive. The proteins have a low pI, which suggests that they would still carry a net negative charge within the environment of the PAGE gel that has a pH of 8.8 and so should still migrate toward the anode. Protein solutions taken from the ferroxidase kinetics experiments were left for 1 hour after the reaction to allow full mineralisation to occur. Insoluble material was removed via centrifugation and the proteins transferred to fresh buffer using the desalting columns as described in section **2.10.3**. Proteins were then concentrated from 0.2 mg/ml to 2.0 mg/ml to allow the loading of 20 μg of protein per lane. Subsequent to electrophoresis and iron detection, positive iron staining was visualised with the appearance of royal blue bands (Figure 3.6B). DpsA and DpsC both gave positive results with iron-loaded proteins. The presence of protein was confirmed by Coomassie staining, which also gave positive results with an overlapping of stains indicating the iron was in fact bound to the proteins. Migration of the proteins was unaffected by the lack of coomassie G-250 sample additive to the gel or within the cathode buffer. Furthermore the presence of an iron core within the proteins did not affect their electrophoretic mobility.

A)



B)

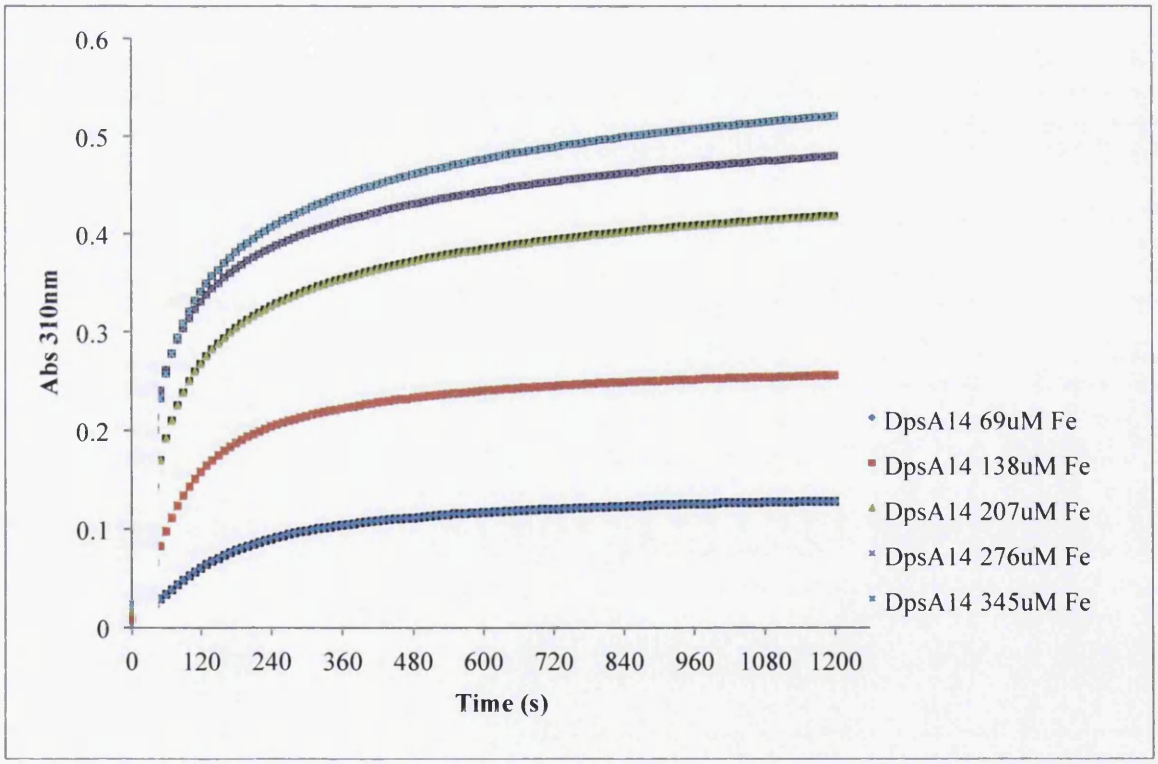
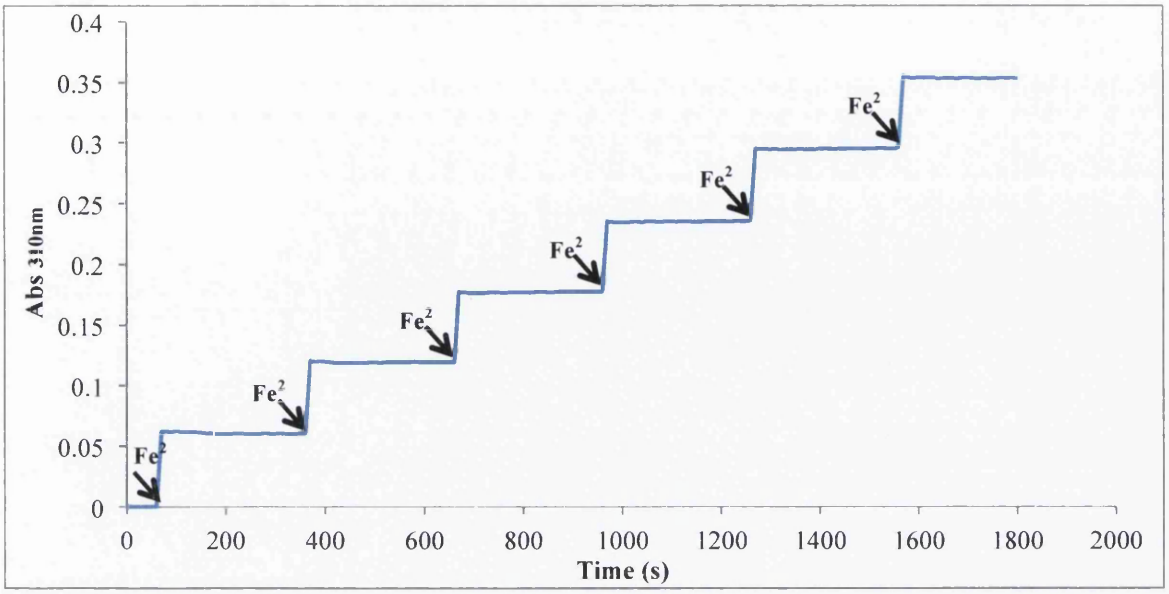


Figure 3.5: (A) Ferroxidase activity of DpsA. 0.2 mg/ml of DpsA oxidising 100 μM ammonium iron (II) sulphate; little oxidation is observed until the addition of 50 μM H_2O_2. Auto-oxidation is observed in parallel. (B) DpsA ferroxidase activity utilising oxygen at various Fe (II) concentrations from 69 μM to 345 μM for each curve, two biological replicates and three technical replicates of each reaction were performed.

A)



B)

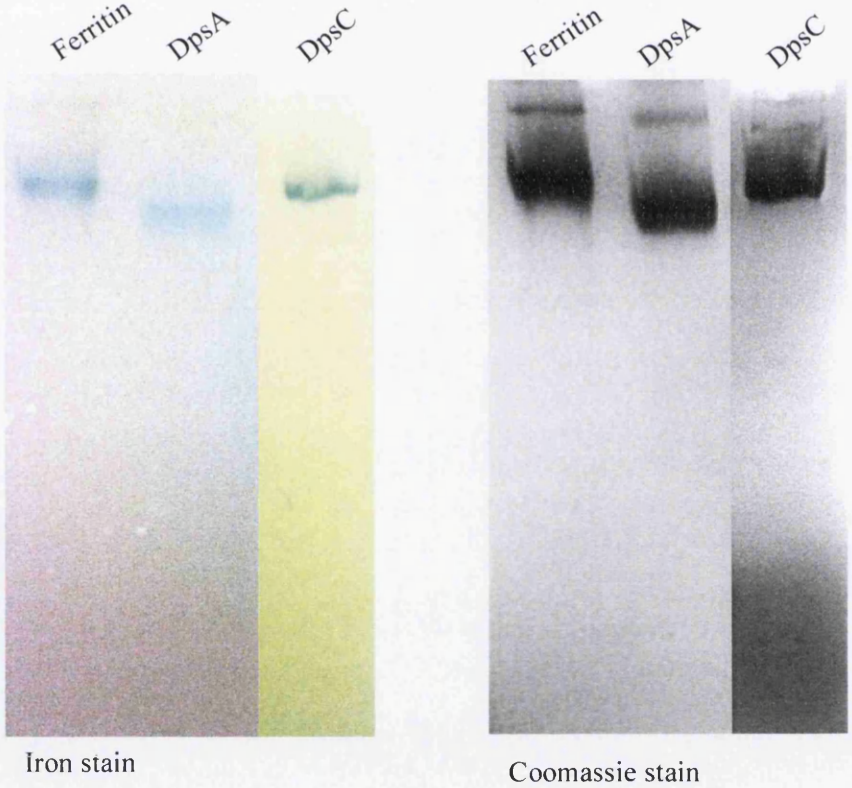


Figure 3.6: Ferroxidase activity and iron binding of DpsC and DpsA. (A) Ferroxidase activity of DpsC utilising hydrogen peroxide. Stepwise addition of hydrogen peroxide (12 μM) followed by addition of Fe (II) (23μM) with six additions of each. Additions of ferrous iron are indicated by the arrows (B) Detection of ferric iron mineralised within DpsA and DpsC proteins resolved by clear native PAGE. Composite gels; lane F – Ferritin control, lane A – Holo DpsA and lane C – Holo DpsC shown as iron staining image (left) and coomassie image (right).

3.3.4 DNA oxidative protection

The functional significance of ferroxidase activity is that the reaction removes the two key elements for Fenton chemistry: ferrous iron and hydrogen peroxide. This nullifies the production of dangerous reactive oxygen species that would otherwise be capable of damaging nucleic acids. A means of monitoring how successful these proteins are at protecting DNA is to perform an *in vitro* DNA Oxidative Protection assay. This assay is carried out by providing all the constituents of Fenton chemistry in a reaction with DNA in the absence and presence of the Dps proteins. Integrity of the DNA was monitored by agarose gel electrophoresis. Whilst DpsB was not included in previous ferroxidase experiment, this assay allows a simple means of discovering if DpsB possesses ferroxidase activity *in vitro*.

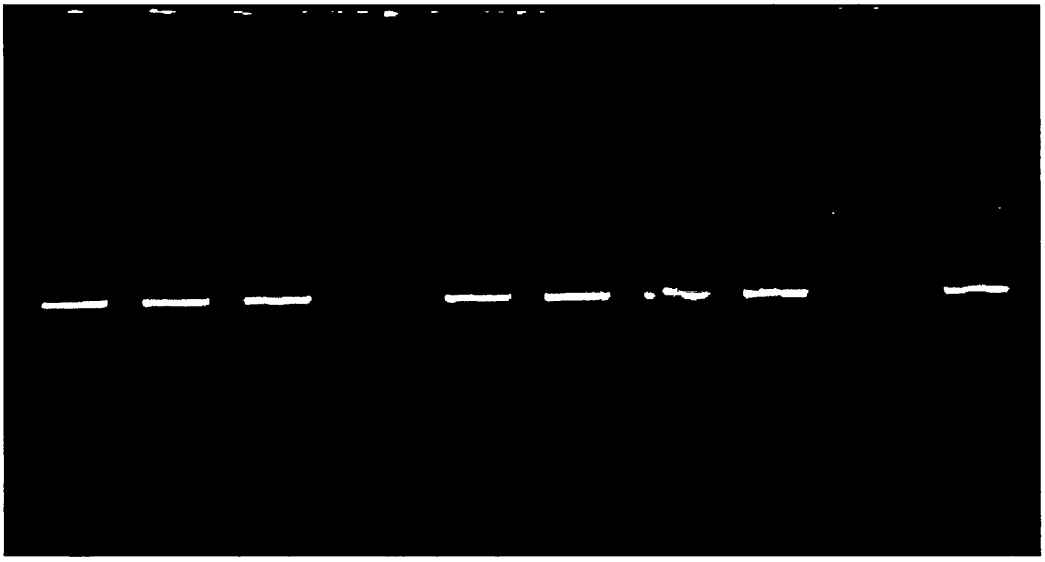
In reactions where only ferrous iron or hydrogen peroxide was incubated with the DNA, the preservation of DNA was observed hence no reactive oxygen species were produced. However the combined effect of the ferrous iron and hydrogen peroxide in the same reaction, incubated with DNA, results in complete degradation of the DNA so that no band is visible in the gel (Figure 3.7A). In the presence of DpsA and DpsC and to a lesser extent DpsB, hydroxyl radical formation is abated by the ferroxidase activity that these proteins possess and as a result the DNA is protected against ROS mediated degradation. The presence of the band in line with that of the control reactions signifies the role of the proteins within the cell, sequestering both the substrates of Fenton chemistry. However, DpsB does not have quite the same ability to protect DNA as a distinct loss of band intensity is seen indicating some DNA degradation. These results also indicate that in the conditions tested, DNA protection is purely the consequence of ferroxidase activity and not the result of a physical interaction between Dps and DNA that would otherwise manifest itself with a shift in band pattern.

3.4 DNA interaction

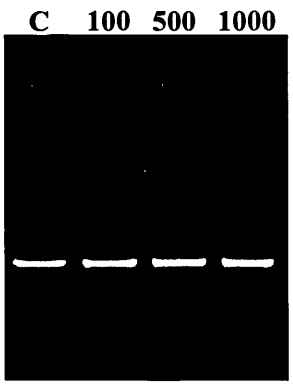
The ability of the *Sc*Dps proteins to directly bind DNA was analysed *in vitro*. An incubation of DNA (pUC18; 0.028 pmol/50 ng) with Dps at DNA/protein molar ratios of 1:100, 1:500 and 1:1000 were incubated at room temperature for 30 minutes in 20 mM Tris/HCl, 200 mM NaCl, and 5% glycerol at pH 7.5 prior to being resolved on a 1% agarose gel. The gel was stained with ethidium bromide post-electrophoresis. No retardation of the DNA was visualised from DpsA, DpsB or DpsC at any of the DNA: protein ratios under the conditions tested (Figure 3.7 B, C and D).

A)

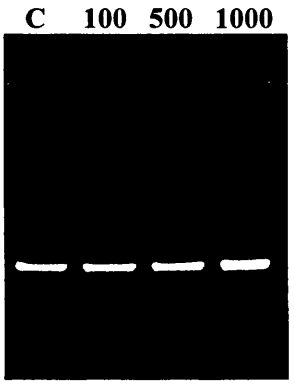
Lane	1	2	3	4	5	6	7	8	9	10
Protein	-	-	-	-	DpsA	DpsB	DpsC	DpsA	DpsB	DpsC
H_2O_2	-	+	-	+	-	-	-	+	+	+
Fe (II)	-	-	+	+	-	-	-	+	+	+



B)



C)



D)

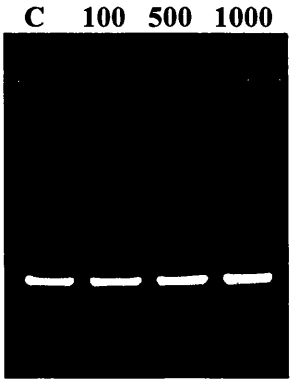


Figure 3.7: *Sc*Dps display DNA oxidative protection but fail to show evidence of DNA binding. (A) Oxidative protection of DNA; DpsA, DpsB and DpsC can protect DNA against reactive oxygen species that form from Fenton chemistry. Reaction conditions are given above the gel. All reaction mixtures contained 0.022 pmol linear pUC18. Lane 4 is the positive control highlighting the complete degradation of DNA in the presence of ferrous iron (50 μM) and excess hydrogen peroxide (5 mM). Electrophoretic mobility shift assay (EMSA) were performed with 0.028 pmol linearised pUC18 (*Hind*III) incubated with Dps proteins at various DNA: Protein molar ratios at 1:100, 1:500 and 1:1000 (labelled above lane). (B) DpsA, (C) DpsB and (D) DpsC showed no DNA mobility retardation when compared to the control (lane C).

3.5 *In vitro* stability of multimeric assemblies

3.5.1 Sensitivity of dodecamers to dissociation by urea

The DpsA and DpsC dodecameric assemblies were subjected to a series of conditions that have denaturing effects on proteins at various levels of protein structure. The first condition tested was urea denaturation. Urea (8 M) is used widely as a mean of disrupting quaternary and tertiary structure and in some cases secondary structure, particularly β-sheets. Helical structures on the other hand can often be maintained (Bennion and Daggett, 2003). Purified DpsA and DpsC were placed in 8 M urea solutions by dissolving urea straight into the protein solution. DpsC contains two cysteine residues. Despite the redox potential of the cellular environment being positive (a reducing environment) the Cys-Cys linkage may still be utilised by DpsC for stabilisation of protein folding and assembly. Therefore to deduce its involvement in dodecameric assembly, the reductant dithiothreitol (DTT) was included at various concentrations in excess (50 mM, 100 mM and 200 mM). After 24-hour incubation at room temperature, the proteins were resolved by native PAGE. The results surprisingly show how DpsC assemblies are resistant to denaturation by urea (Figure 3.8A). The preservation of quaternary structure is emphasized by the persistence of the high oligomeric band corresponding to a dodecamer. The inclusion of DTT had a mild effect on the lower oligomeric species, causing a slight shift in the band to a smaller size. However the band shift is not large enough to be the result of the dissociation of monomers from each other, but is potentially an artificial effect of the urea unfolding process, which reveals cysteines that are normally buried, and solvent inaccessible. A partially unfolded monomer can adopt an alternative tertiary structure that allows the thiol groups of cysteine residues to become close enough to form disulphide bridges. These can then be reduced with the addition of DTT. DpsA assemblies behaved more normally. DpsA quaternary structure is disrupted by urea causing the dodecameric assembly to break apart entirely (Figure 3.8A). There is also no effect caused by DTT since DpsA has no cysteine residues within its sequence.

Since 8 M urea was capable of dissociating the dodecamer of DpsA, a gradient of urea concentrations from 0 to 8 M at 1 M intervals was created in order to find the minimum concentration required to disrupt dodecameric assembly. Figure 3.8B demonstrates that with increasing concentrations of urea, the abundance of the lower oligomeric species intensifies. Between the concentrations of 1 M and 3 M, urea has very little effect on assembly. At 5 M urea, almost all of the dodecamer has dissociated, and from 6 M urea upwards only the lower oligomeric bands exist.

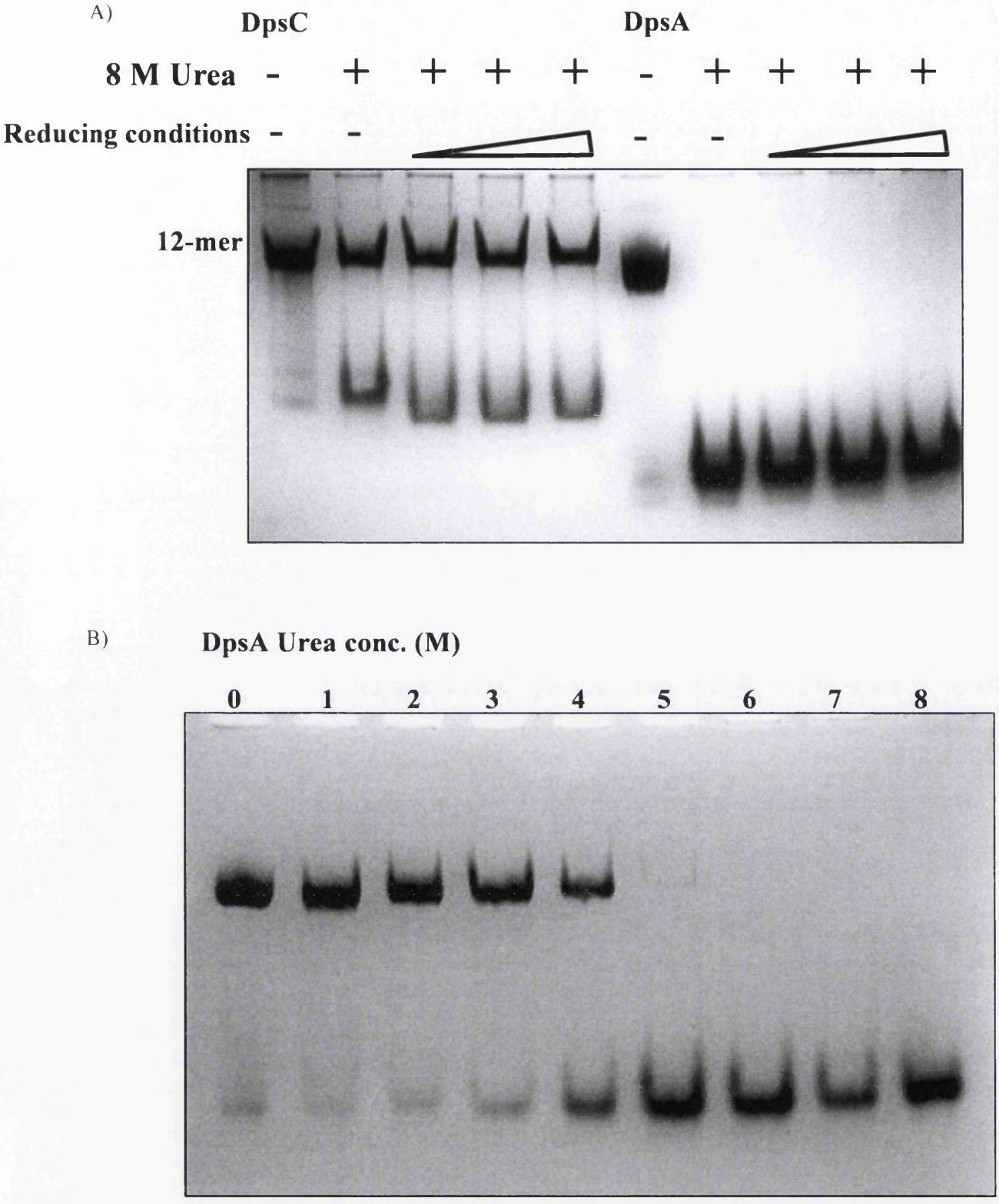


Figure 3.8: DpsC and DpsA susceptibility to denaturation with 8 M urea (A). DpsC remains a dodecamer and is unaffected by 8 M urea even in reducing conditions (50 mM, 100 mM and 200 mM DTT). (B) DpsA can be dissociated with urea at a concentration of 5 M without reducing conditions being required. At lower urea concentrations DpsA dodecameric assembly is unaffected.

3.5.2 Influence of pH on dodecameric stability

The stability of the Dps dodecamers was tested across a pH range that covers both acidic and alkaline conditions. After transferring the proteins to the relevant pH solution via buffer exchange columns, the proteins were incubated for 24 hours prior to resolving their oligomeric state by BN-PAGE. DpsA started to dissociate from dodecameric assembly at pH 5.0. At pH 4.0 the absence of the dodecameric band but presence of a lower oligomeric state species confirms that the assembly is disrupted at this pH (Figure 3.9A). There is also the presence of protein still remaining in the wells of the gel after electrophoresis which can be attributed to the self-aggregation phenomenon reported to occur for Dps proteins at acidic pH levels (Ceci et al., 2004). This self-aggregation is thought to result from the properties of the N-terminal tails of the Dps. Another plausible explanation for the residual protein left in the well is the relationship between protein charge and solution pH. When proteins are in an aqueous environment where the pH matches the isoelectric point (pI), the proteins carry no charge; the charge-based interactions that occur with water cannot form. This favours a tendency for protein-protein interactions to take place, which usually manifests itself in protein aggregation.

In acidic conditions DpsC shows no signs of dissociation or aggregation. Unlike DpsA, there are no signs of any low oligomeric state bands present at pH 5.0 or 4.0 but the protein would appear to be a highly stable dodecamer.

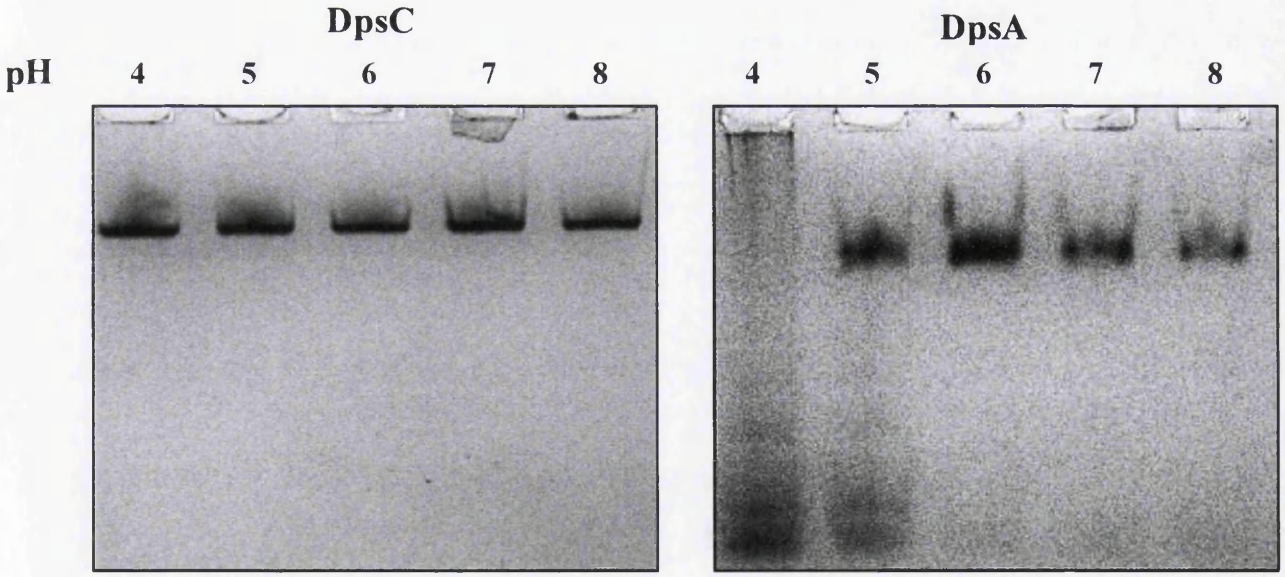


Figure 3.9: pH stability of DpsC and DpsA between pH 4.0 and 8.0. DpsC remains as a dodecamer at lower pH levels while DpsA starts to dissociate at pH 5.0 and no observable dodecameric assemblies remain at pH 4.0 suggesting dissociation to a low oligomeric assembly.

3.6 Summary of Recombinant Expression and Biochemical characterisation of *S. coelicolor* Dps proteins results

This first results chapter has dealt with the basic exploration of the Dps proteins from *S. coelicolor*. The objectives were to give a brief insight into how the proteins behave and function both *in vivo* and *in vitro* and to generate data that allows comparison not only between the three proteins but also between other members of the Dps protein family. A basic summary of the results is given below;

- Over expression of DpsA *in vivo* does not alter its oligomeric state compared to its expression after heat stress and osmotic shock
- DpsA and DpsC assemble into dodecamers *in vivo* and *in vitro*
- DpsB fails to assemble into dodecamers both *in vivo* and *in vitro*
- The N-terminal tail of DpsA is implicated in oligomeric assembly since placing a poly-His tag at the N-terminal effects dodecameric assembly
- At neutral pH, DpsA is capable of catalysing the oxidation of Fe (II) to Fe (III) in the presence of hydrogen peroxide or oxygen
- DpsA ferroxidation is more efficient using H_2O_2 as an oxidant than O_2
- DpsC requires H_2O_2 for efficient oxidation but requires a step wise addition of iron to prevent protein precipitation
- Both DpsA and DpsC are capable of binding ferric oxide, depositing the mineralised iron oxide within the cavities of their spherical cages
- The three *Sc*Dps proteins can protect DNA against oxidative damage that would occur via Fenton chemistry produced ROS.
- Under the conditions tested, no DNA binding ability was detected for any of the proteins
- DpsA assembly can be dissociated by 8 M urea and acidic pH (pH 4.0)
- DpsC maintains a highly stable and resistant multimeric assembly in unfavourable environments; denaturing conditions and low pH

These results provide a number of questions that will be explored in the following chapters. The difference in stability between DpsA and DpsC can be explained by resolving the crystal structures of the proteins. Furthermore the role of the N-tail of DpsA will be explored in detail using the structure and by creating mutant proteins. These subjects will be the foci of the following chapters.

Chapter 4

Resolution and Analysis of the Crystal Structures *Sc*DpsA and *Sc*DpsC

4.1 Introduction

Protein crystallography has had a profound effect on molecular biology, particularly the field of structural biology and biochemistry. The elucidation of vast numbers of protein structures from far ranging fields of biology has been a prerequisite for deducing and resolving the regulation and function of macromolecules. This method, together with other *in vitro* molecular biology data, allows the full characterisation of proteins to be carried out down to the atomic level, allowing the structure to provide further insight into how these proteins perform their roles The elucidation and deposition of a number of Dps-related structures in to the PDB database allows for comparisons between proteins which are structurally and functionally similar or diverse in order to relate structure to function.

DpsA and DpsC proteins were sent for out of house analysis (Durham University Chemistry Department, in care of Dr Ehmke Pohl) to determine their crystal structure in their apo state. This chapter deals with the analysis of the crystal structures of DpsA and DpsC from structural validation through to exploration of macromolecular interfaces and surfaces presenting interesting features that contribute to the functions of Dps. This also provides a base for comparison with other Dps structures and reveals extensive differences between DpsA and DpsC.

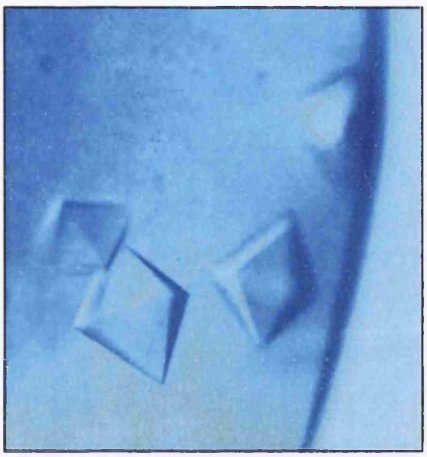
4.2 Crystallisation conditions

The crystallisation process is very much empirical and conditions required for crystal formation are found through a trial and error approach. However, there are commercially available screening systems that can be used that provide a range of parameters to be screened. In this case crystallisation conditions were tested using a JCSG-plus sparse-matrix screen, which provides a range of environments for crystal growth based on previously successful outcomes. Pact Premier was simultaneously used to trial conditions relating to pH, anion and cation composition in a PEG/ion screen environment. Multiple conditions within the Pact Premier kit were found in these coarse screenings to produce good crystal formations, these conditions are provided in Table 4.1, however conditions of 0.2 M sodium chloride, 0.1 M sodium cacodylate, 2.0 M ammonium sulfate pH 6.5 for DpsA (E2) and 0.1 M Bis Tris pH 5.5 with 3.0 M sodium chloride for DpsC (G12) from the JCSG-plus sparse-matrix screen were found to be optimum.

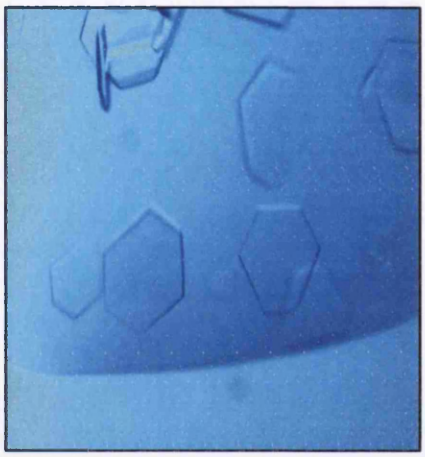
Table 4.1.Coarse conditions for crystal growth of DpsA and DpsC

Buffer/Salt	Buffer/Salt	pH	Precipitant
0.2 M Di-ammonium hydrogen citrate	None	5.0	20 % PEG 3350
0.1 M SPG buffer	None	6.0	25 % PEG 1500
0.2 M Calcium Chloride	0.1 M bis-Tris	5.5	45 % MPD
0.2 M Ammonium Chloride	0.1 M bis-Tris	5.0	25 % PEG 3350
0.2 M Magnesium Chloride	0.1 M MES	6.0	20 % PEG 6000
0.1 M PCPT Buffer	None	5.0	25 % PEG 1500

A)



B)



C)

A

MTHDLTPKYTVPGIEREAAGRLIGVLRLRLHALNDLHLTLKHVHWNVVGPH

B BC

FIAVHEMIDPQVDQVRDMADDVAERIAALGGVAQGTPGALVAERKWDDYS

C D

IGRADAIAHLGALDVVYTGVVEGMRAAVEEAGKIDPATEDLLIGQLRDLEQF

QWFVRAHLESAGGALATGSATSETEAAARGAEVADPNSSSVDKLAAALEHHH

HHH

Figure 4.1: Growing crystals of Dps proteins; DpsA (A), DpsC (B). Images were taken at 15 times magnification. (C) DpsA protein sequence. Highlighted in grey are the helical regions which are labelled above A through to D. Residues not resolved in the crystal structure are underlined with the linker region and His tag double underlined. Bold red residues are those that create the conserved ferroxidase centres.

4.3 Structural validation

4.3.1 Structure validation of DpsA

The crystals of DpsA that were studied diffracted to 1.78 Å on a IO4 beamline (Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxford, UK) and yielded one monomer per asymmetric unit with amino acids corresponding to 4-166 of the 187 amino acids present in the native DpsA protein sequence. However, the total protein sequence allowing for the histidine tag and a linker region between the protein's C-terminal and the His tag accounts for 208 amino acids thus a sequence of 39 amino acids at the C-terminal was regarded as being disordered in the crystal structures. If consideration is paid to the native DpsA sequence, only 18 residues of the C-terminal tail are missing from the structure with 3 amino acid residues missing from the N-terminal tail. This inability to model the terminal residues has been previously documented with other Dps protein structures such as the *Deinococcus radiodurans* Dps, in which 163 out of 207 residues were modelled (Romão et al., 2006). In this case, it is thought the structural flexibility of the tail is associated with functions such as DNA interaction (Grant et al., 1998).

The atomic co-ordinates, as a PDB file, were submitted to the Protein Structure and Validation Server version 4 (http://services.mbi.ucla.edu/SAVES/) which can run a series of programs for validation of 3D structures. The ERRAT program used to detect errors in model building by analysing the pattern of non-bonded interactions (Colovos and Yeates, 1993) yielded an overall quality factor of 99.359 %, which is above the 95 % rejection limit (Appendix 8). Furthermore a Ramachandran plot (created using http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) is well behaved, showing that 98.8 % of all residues are in the favoured region with just 1.2 % of amino acids in the allowed region (Appendix 9). This means that there is very little steric hindrance found within the structure. If each atom is treated as a sphere with dimensions which correspond to their Van der Waals radii, the high percentage of residues within the favoured region have phi (N-Ca torsion angle) and psi (Ca-C torsion angle) angles that do not cause radii on side chains to collide or clash with radii in the main chain.

4.3.2 Structure validation of DpsC

Crystals of DpsC yielded a full dodecamer per asymmetric unit and diffracted down to 1.78 Å on an IO2 beamline (Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxford, UK). As with DpsA, the recombinant DpsC sequence is larger than the native sequence; combining the linker region and the His tag, recombinant DpsC amounts to 221 amino acids in length. Based on secondary structure prediction, DpsC possesses a very long N-

terminal tail, 44 residues long, but only 17 amino acids were visualised in the structure, with the other 27 amino acids disordered in the crystal structure. Yet this is still the joint highest number of residues resolved of an N-terminal tail of the Dps structures within the PDB, equal with the structure of the *Thermosynechococcus elongatus* Dps (PDB 2VXX). On the contrary to the N-terminal tail, DpsC is predicted to have an 8 residue C-terminal tail and the structure resolved all eight residues from the end of the D helix onward. Furthermore the linker region and His tag of 21 residues in total could not be resolved. The total number of residues found within the crystal structure was 173. This means a total of 48 amino acids were not resolved.

The Protein Structure and Validation Server was used to validate the DpsC 3D structural co-ordinates. The ERRAT program yielded an overall quality factor of 98.919 % (Appendix 10), again above the 95% cut off limit. The Ramachandran plot for DpsC (Appendix 11) shows that 98.8 % of all residues are within the favoured region and as with DpsA just 1.2 % in the allowed region.

MSSPKPKPPSSAEHRSDGSQPWLHQKGRTIQEFGTVKQFPVALTMDTRLYSC

A B

QRLNKVLADTRILHDLYKKYHWLMRGATFYQLHLLLDKHAGEQLELIDTV

BC C

AERVQTLGGVAVGDPRHVAEITTVPRPPDGVEEVPSMLSRLLEAHELILTEC

D

HDAAARTQEYGDDGTNDLLVSEVLRTNELQAWFVAEHLVDTPLVHADPNSS

SVDKLAAALEHHHHHH

Figure 4.1.1: The annotated protein sequence of DpsC. Highlighted in grey are the helical regions which are labelled above A through to D. Residues not resolved in the crystal structure are underlined with the linker region and His tag double underlined. Bold red residues are those that create the conserved ferroxidase centres.

4.4 Overall structure

4.4.1 Monomeric structure (tertiary) of DpsA - The Stereotypical Four Helix Bundle

As with other Dps and Dps-like structures present in the PDB database, the DpsA monomer shown in Figure 4.2A is a square classed four helix bundle with helices denoted as A, B, C and D which correspond to residues 16-46 (A), 52-80 (B), 107-134 (C) and 137-161 (D). There is an additional small BC helix which is locate perpendicular to the cylindrical four helix bundle in the long loop between the B and C helices that transverses the entire length of the helix bundle. This B/C helix is composed of residues 88-94. There are non-helical regions located at the N-terminal (12 residues long) and C-terminal (8 residues long) regions which make up the features known as "tails", in addition to loops located between helices. Helices account for 120 residues out of 167, this equates to 71.8 % of the protein.

Each helix has a slightly different length as the number of residues forming each helix differs. The distances can be measured within UCFS Chimera 1.7rc using the atomic distance tool between alpha carbons. The A helix, composed of 31 amino acids, has a length of 49.198 Å between Asn-16 and Trp-45. The B helix, composed of 29 amino acids, and has a length of 45.326 Å between Ile-53 and Leu-80. The BC helix of 7 amino acids is 13.602 Å between Pro-88 and Glu-94. The C helix is 28 residues long and measures 43.515 Å between Ala-107and Lys-134. Finally the D helix is 25 amino acids long and is 39.667 Å between Pro-137 and His-160. The distance between the alpha-carbon of the residue Glu-15 at the N-terminal region and the alpha-carbon of residue Pro-50 at the opposite end of the monomer is approximately 59.440 Å. This is the largest distance between two residues in the whole monomer including the loops and the N-and C-terminal tail extensions. The actual four helix bundle has a length of approximately 54.337 Å which is between the alpha-carbon of Arg-16 at the N-terminal region of the A helix and the alpha-carbon of Asp-106 at the C-terminal region of the C helix. This distance is longer than the A helix, the longest helix, because the bundle is not strictly square. Within the subunit there are a number of intra- and inter-helical interactions that help to maintain the structure. A total of 329 hydrogen bonds can possibly be formed within a monomer including intra-helical hydrogen bonds that are vital to the formation and stability of alpha helices (Figure 4.2C). Furthermore ten salt bridges are also predicted to be found within the monomer and are reported in table 4.2

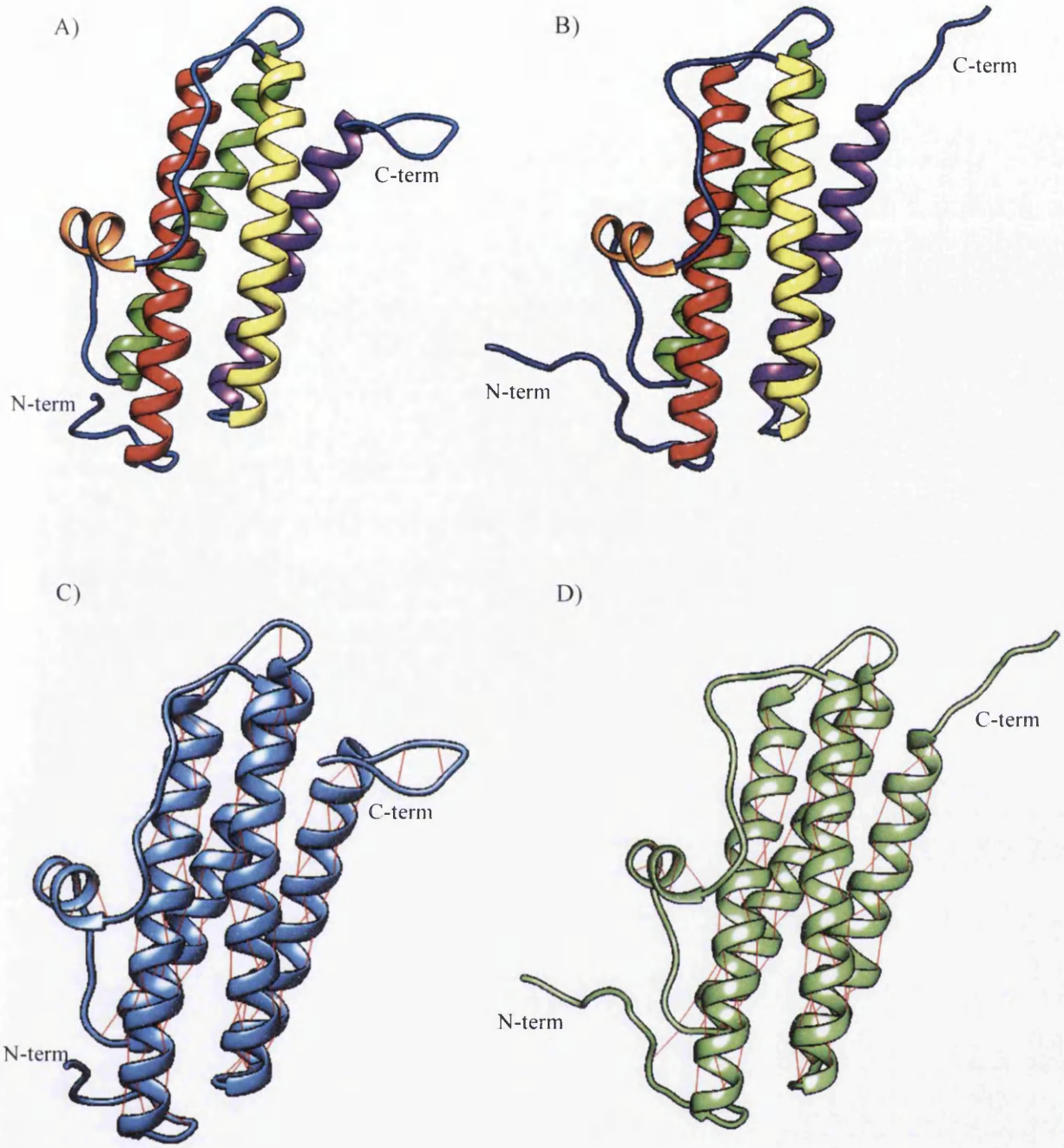


Figure 4.2: Monomeric structures of DpsA and DpsC. Ribbon diagram of recombinant DpsA crystal structure showing how secondary structure elements such as the alpha helices form the tertiary structure four helix bundle. (A) DpsA, the entire four helix bundle (B) DpsC, the entire four helix bundle. The helices are color coded; A-red, B-green, B/C, orange, C-yellow and D-purple. (C) DpsA four helix bundle with 329 hydrogen bonds indicated with red lines. (D) DpsC four helix bundle with 193 hydrogen bonds indicated with red lines. Structures do not contain hydrogens, so bonds are draw between donor and aceptor groups. Images were created using UCFS Chimera 1.7rc.

Table 4.2: Residues predicted to participate in salt bridge formation in DpsA monomer

Residue 1	Residue 2	Distance Å
NH2 ARG A 21	OE1 GLU A 17	3.37
NH2 ARG A 27	OE1 GLU A 94	3.56
NE2 HIS A 31	OE1 GLU A 94	2.86
NH2 ARG A 76	OD1 ASP A 136	3.95
NH2 ARG A 76	OD2 ASP A 136	2.78
NH1 ARG A 95	OD1 ASP A 99	3.08
NH2 ARG A 95	OD2 ASP A 35	2.85
NH2 ARG A 95	OD1 ASP A 99	3.00
NH2 ARG A 126	OE1 GLU A 151	3.42
NH2 ARG A 158	OD1 ASP A 115	3.88

4.4.2 Monomeric structure (tertiary) of DpsC

DpsC also yields a monomer, which obeys the structural characteristics of Dps proteins. A four-helix bundle composed of A, B, C, D and the small BC helix is observed Figure 4.2. Helix A is composed of 31 amino acids and is 44.403 Å in length from Met-45 to Leu-75. Helix B is 28 residues in length and measures 42.005 Å between Phe-81 and Gln-107. The BC helix runs from residues 117-123 and is 17.298 Å between Arg-118 and Ile-123. The C helix is 44.341 Å at the longest distance between Val-136 and Glu-163 which are the terminal residues of the helix thus this helix is composed of 28 residues, with the D helix 26 residues long, starting at Asp-167 and finishing at Leu-192. This helix measures 41.133 Å between Asp-167 and His-191. The largest distance running the entire monomer length without measuring diagonally is 68.540 Å between Gln-38 and His-199. This includes the N and C terminal tails. However the four-helix bundle measures 55.231 Å between Met-45 and Ala-79. Of the 173 amino acids present within the crystal structure, 120 of them form helical structures. This equates to 69.36 % of the structure dedicated to helices. There are 193 intra-monomer hydrogen bonds that lend themselves toward stabilising the helical structures and thus the monomer on the whole Figure 4.2. In addition to these hydrogen bonds, 15 salt bridges were also predicted to form between the residues shown in the table 4.2, an increased value compared to DpsA and as previously described is a mode employed for increased thermal stability in proteins.

Table 4.2.1: Residues predicted to participate in salt bridge formation in DpsC monomer

Residue 1	Residue 2	Distance Å
NH1 ARG-54	OD2 ASP A 156	2.80
NH2 ARG A 54	OE1 GLU A 153	3.33
NH2 ARG A 54	OD2 ASP A 156	2.93
NZ LYS A 57	OD1 ASP A 61	3.32
NZ LYS A 57	OD2 ASP A 61	2.62
NE2 HIS A 66	OD1 ASP A 67	3.85
NZ LYS A 70	OD1 ASP A 89	2.84
NE2 HIS A 73	OD1 ASP A 89	3.85
NZ LYS A 90	OE1 GLU A 94	2.90
NH2 ARG A 105	OD1 ASP A 166	2.86
NH2 ARG A 118	OD1 ASP A 116	3.57
NH2 ARG A 118	OD2 ASP A 116	2.82
NH2 ARG A 128	OD1 ASP A 131	2.80
NH2 ARG A 128	OD2 ASP A 131	3.61
NH1 ARG A 160	OE1 GLU A 163	3.41

4.4.3 Multimeric structure (quaternary) of DpsA

The quaternary structure of DpsA is a dodecamer with 23-point symmetry (Figure 4.3). Composed of 12 subunits this complex multimeric structure is stereotypical of other members of the Dps family, for example; *Mycobacterium smegmatis* (Roy et al., 2004), *Agrobacterium tumefaciens* (Ceci et al., 2003) and *Salmonella enterica* (Miyamoto et al., 2011b), which all return as the top three hits respectively, when searching the PDB database with the DpsA protein sequence as a query. It was thus appropriate to use the *M. smegmatis* Dps structure (1UVH) on which to chain-trace DpsA. The dodecamer forms a hollow spherical protein cage with an external diameter of approximately 87.202 Å from alpha-carbon of Lys-96 to the alpha-carbon of Lys-96 of the subunit on the opposite side of the dodecamer. These residues are situated in the loop between the BC helix and the C helix. This value is congruent with measurements reported from other Dps proteins (Ren et al., 2003). An internal diameter of around 36.234 Å, is slightly smaller than reports from other Dps structures, however as the cavity is less circular there is variation in diameter depending on the atoms specified and since few reports specify the atoms used to measure distances it is hard to compare like for like.

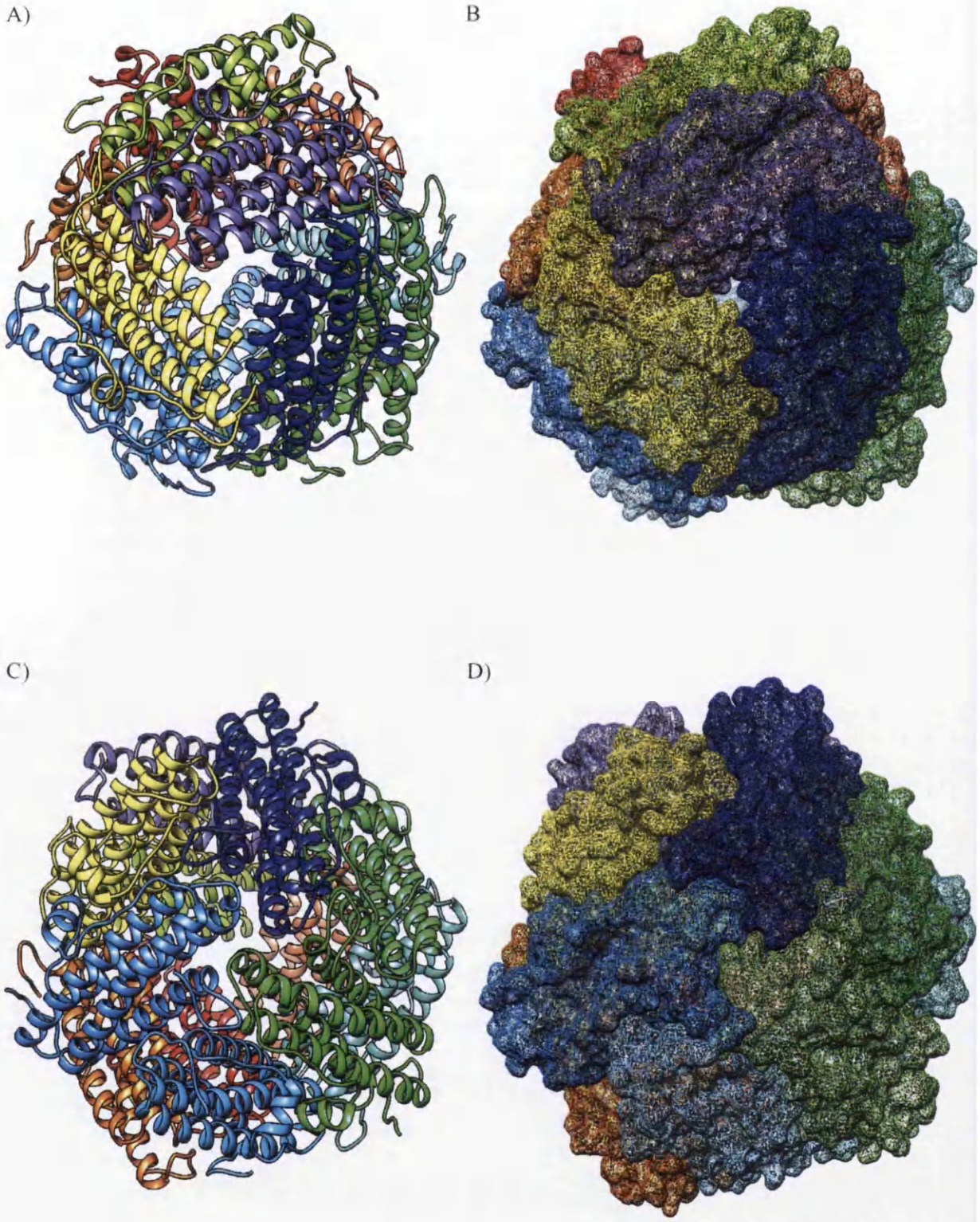


Figure 4.3: Crystal structure models of the DpsA dodecamer. (A-B) view from the N-terminal ferritin-like pore; ribbon and mesh surface, (C-D) view from the C-terminal Dps-like pore; ribbon and mesh surface. Each subunit is tinted in a different colour highlighting the symmetry that exists in the dodecamer. Images were created using UCFS Chimera 1.7rc

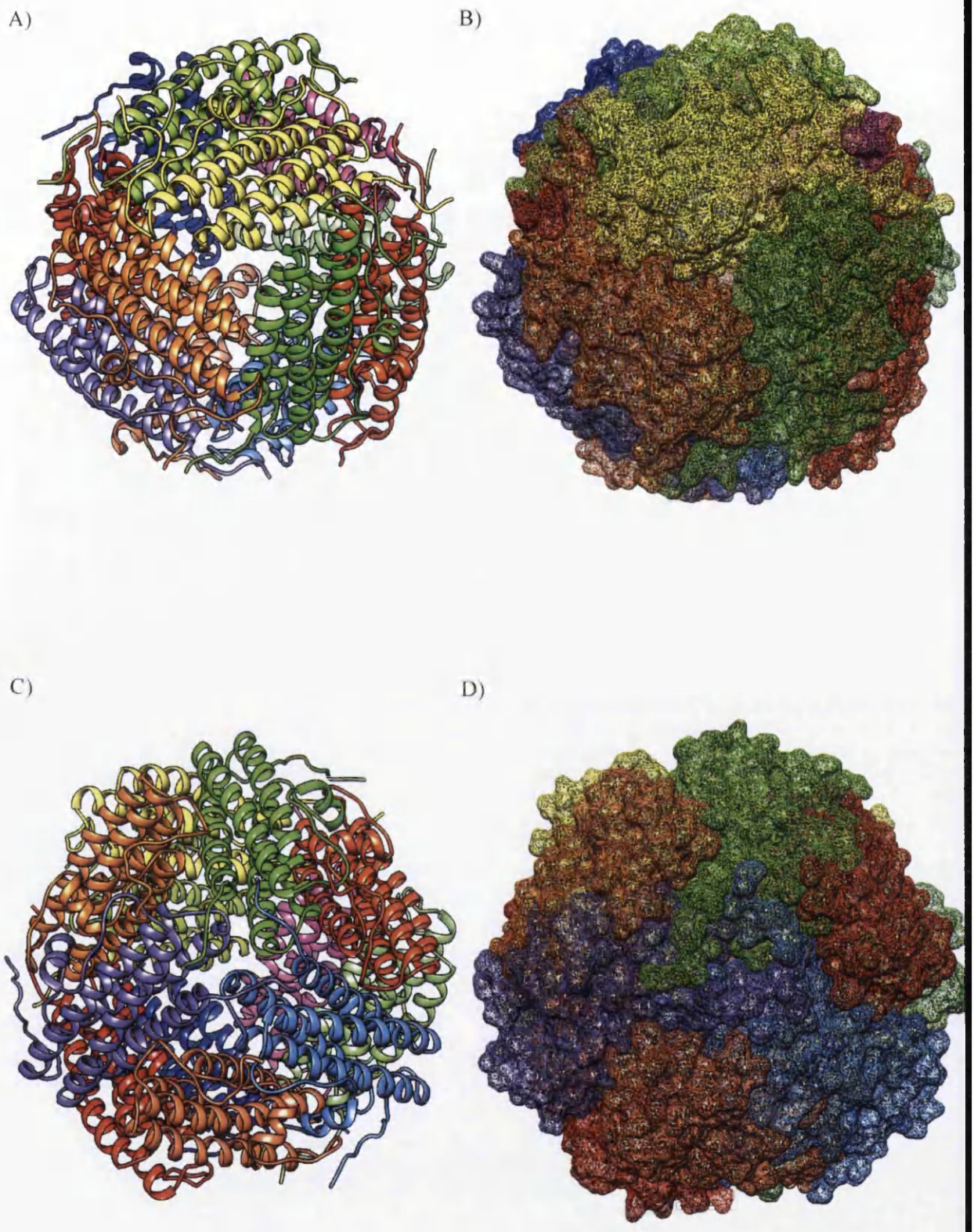


Figure 4.4: Crystal structure models of the DpsC dodecamer. (A-B) view from the N-terminal ferritin-like pore; ribbon and mesh surface, (C-D) view from the C-terminal Dps-like pore; ribbon and mesh surface. Each subunit is tinted in a different colour highlighting the symmetry that exists in the dodecamer. Images were created using UCFS Chimera 1.7rc

4.4.4 Multimeric (quaternary) structure of DpsC

The quaternary structure of DpsC also maintains the conventional spherical dodecameric assembly with 23-point symmetry (Figure 4.4). As there are very few proteins that have high sequence similarity with DpsC, let alone with a resolved crystal structure, DpsA was used for chain tracing of the DpsC structure. The dodecamer's diameter, measured from residue Glu-122 on one monomer to Glu-122 on the monomer on the opposite side of the structure, gave a distance of 86.685 Å. These residues are located on the BC helix of the dodecamer and protrude outwards towards the solvent. Being a spherical cage, DpsC also has an internal cavity which has a diameter of approximately 34.491 Å. This is congruent with a slightly smaller external diameter when compared to DpsA. It is also worth noting that measurements are based on the atomic models, and this does not take into account the molecular surfaces.

4.4.5 Structural Superposition of DpsA and DpsC

Superpositioning of proteins allows their structures to be super-imposed pairwise by first aligning their sequences followed by the fitting of alpha carbons of residues in each column of the alignment. This allows the similarity of protein folding to be examined when sequence divergence may suggest a dissimilar protein-folding pattern; although actual folding and residue side chain conformations may in fact be highly homologous. All Dps protein monomers available from the PDB were super positioned on top of DpsA and DpsC monomeric structures and the Root Mean Square Deviation (RMSD) reported. The tables are provided as Appendices 12 and 13, while structural superposition of DpsA and DpsC is shown in Figure 4.5. The RMSD of DpsC superimposed on DpsA is 1.047 Å between 136 atom pairs. DpsA was found to be most similar to the *M. smegmatis* Dps (PDB ID 1VEI) (Roy et al., 2004) with a low RMSD of 0.681 Å between 160 atom pairs. Conversely the Dps of *B. anthracis*, (PDB ID 1JI5) (Papinutto et al., 2002), gives a high RMSD of 1.231 Å between 128 atoms when compared to DpsA.

DpsC has a structure that was most similar to the Dps from the thermophilic cyanobacterium *T. elongatus* (PDB ID 2C41) (Franceschini et al., 2006) with a very low RMSD of 0.761 Å between 130 atom pairs. The Dps from *D. radiodurans* (PDB ID 2C2F) (Romão et al., 2006) was the most distinct from DpsC with a RMSD between 120 atoms of 1.158 .

Most notable from these superpositions are the placement of the loop and tail structures which appear to be very similar within the structures that have low RMSD values and very different in comparisons that have high RMSD values. Thus these represent the main regions of difference between the Dps proteins examined.

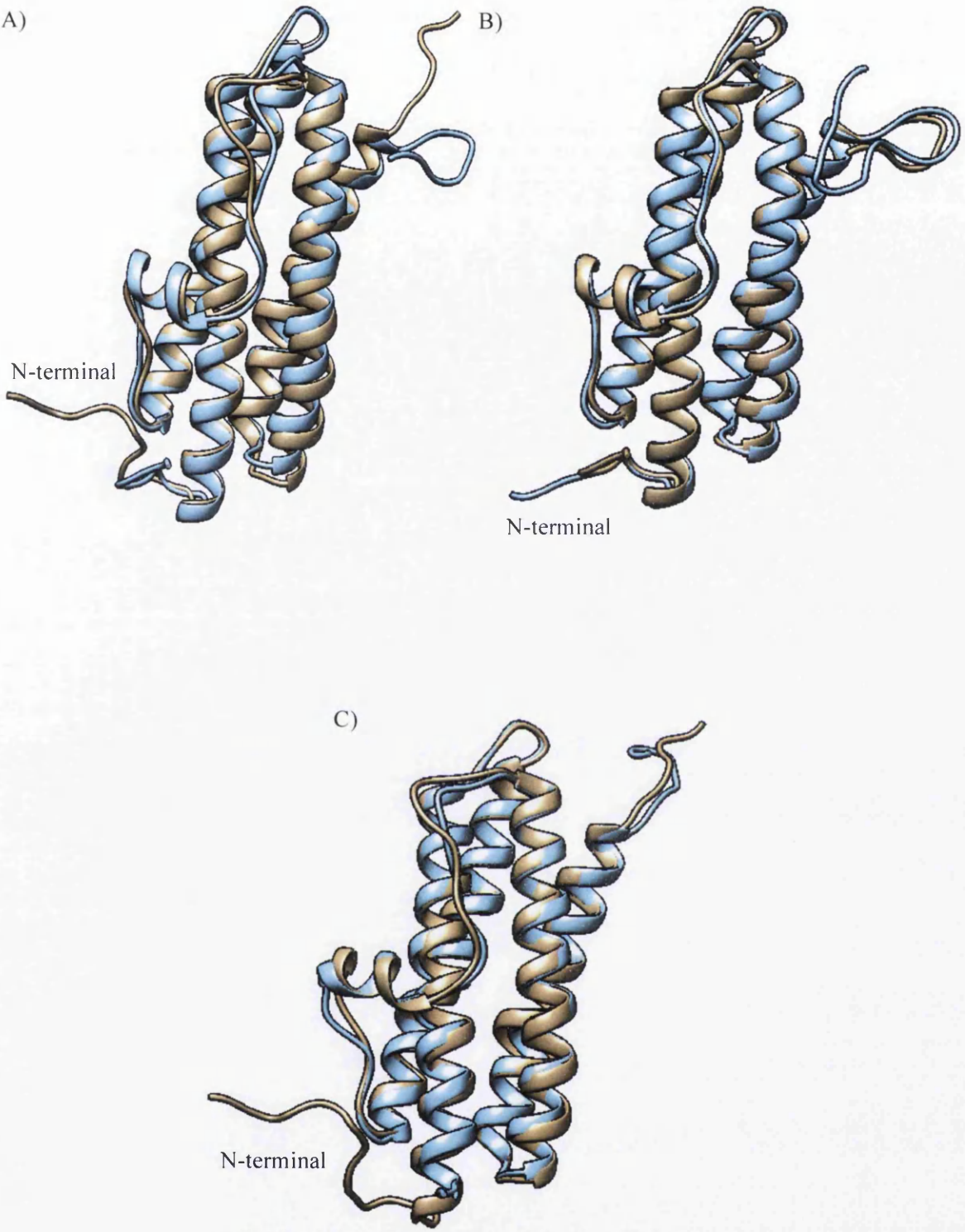


Figure 4.5: Structural Superposition Alignment of Dps monomers. (A) Superposition of DpsA (blue) and DpsC (brown) with an RMSD of 1.047 between 136 atom pairs (B) Superposition of DpsA (brown) and *M. smegmatis* Dps (1VEI) (blue) RMSD of 0.681 between 160 atom pairs (C) Superposition of DpsC (brown) and *T. elongatus* Dps (2C41) (blue) RMSD of 0.761 between 130 atom pairs.

4.5 Identification and analysis of structural and chemical features of *Sc*Dps

It is possible to see from the crystal structure that the *Sc*Dps proteins retain the structural characteristics conserved within the Dps family; from the helical secondary structures, to the four helix bundle tertiary structure and onto the dodecameric quaternary structure. However, it is also possible to predict, study and compare the functionally relevant regions of these proteins to each other and Dps from other organisms.

4.5.1 Proteins, Interface, Structure and Assemblies Analysis (PISA)

The PDB files for DpsA and DpsC were submitted to the PDBePISA interactive tool for analysis of macromolecular interactions and interfaces (Krissinel and Henrick, 2007), available at http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html. Atomic co-ordinates were submitted with the intention of identifying interfaces present in the multimeric assemblies and residues involved in these interfaces. These interfacing regions are important, inter-monomer interactions that give rise to the dodecamer. It is thus the chemical properties and strength of these individual interactions that govern the strength of the dodecamer on the whole.

Three major types of interfaces were detected when analysing the DpsA and DpsC structural co-ordinates; dimer interactions at two-fold symmetry axes in addition to two types of trimeric interactions that maintain three-fold rotation axes formed by non-equivalent N-terminal interactions and C-terminal interactions; a summary of the interfaces is given in table 4.3 with a legend to this table provided as table 4.4. The interface types predicted by this analysis are highly conserved within the Dps family of proteins; the dimer interface, N-terminal trimer "ferritin-like" interface and the C-terminal trimer "Dps-like" interface, however the exact chemical nature differs.

Table 4.3: A summary of interface characteristics found in DpsA and DpsC dodecamers.

	Interface type	Interface area, $Å^2$	$\Delta^i G$ kcal/mol	$\Delta^i G$ P-value	N_{HB}	N_{SB}	CSS
DpsA	Dimer interface	1310.3	-7.9	0.552	14	8	0.660
	C-terminal trimer	512.0	-7.1	0.260	2	2	0.697
	N-terminal trimer	742.0	-3.5	0.552	8	7	0.295
DpsC	Dimer interface	1870.5	-12.8	0.699	27	27	0.525
	C-terminal trimer	820.4	-14.2	0.135	5	0	0.706
	N-terminal trimer	735.0	-4.9	0.637	8	9	0.593

Table 4.4: Proteins, Interface, Structure and Assemblies Analysis (PISA) output descriptions

Interface area, A^2	**Interface area** in $Å^2$, calculated as difference in total accessible surface areas of isolated and interfacing structures divided by two.
$\Delta^i G$, kcal/mol	$\Delta^i G$ indicates the solvation free energy gain upon formation of the interface, in kcal/M. The value is calculated as difference in total solvation energies of isolated and interfacing structures. Negative $\Delta^i G$ corresponds to hydrophobic interfaces, or positive protein affinity. This value does not include the effect of satisfied hydrogen bonds and salt bridges across the interface
$\Delta^i G$, P-value	**$\Delta^i G$ P-value** indicates the P-value of the observed solvation free energy gain. P-value measures the probability of getting a lower, then obtained, $\Delta^i G$, if interface atoms are picked randomly from protein surface such as to amount to the observed interface area. P-value is a measure of interface specificity, showing how surprising, in energy terms, the interface is. For example, **P=0.5** means that the interface is not "surprising" at all with $\Delta^i G$ equal to an average value for given structures. **P>0.5** means that the interface is less hydrophobic then it could be, therefore the interface is likely to be an artefact of crystal packing. **P<0.5** indicates interfaces with surprising (higher than would-be-average for given structures) hydrophobicity, implying that the interface surface can be interaction-specific. The limiting case of **P=0** means that no other interface of the observed area may have a lower $\Delta^i G$, therefore such interface is a truly unique spot on protein surface
N_{HB}	N_{HB} indicates the number of potential hydrogen bonds across the interface. Each hydrogen bond contributes about 0.5 kcal/mol into the free energy of protein binding (exact value depends on the calibration procedure and may change with version number).
N_{SB}	N_{SB} indicates the number of potential salt bridges across the interface. Each salt bridge contributes about 0.3 kcal/mol into the free energy of protein binding (exact value depends on the calibration procedure and may change with version number)
CSS	**CSS** stands for the Complexation Significance Score, which indicates how significant for assembly formation the interface is. The score is defined as a maximal fraction of the total free energy of binding that belongs to the interface in stable assemblies. For example, if the only stable assembly found is a tetramer with 2 interfaces #1 with ΔG=-10 kcal/M and 2 interfaces #2 with ΔG=-5 kcal/M, then CSS for interface #1 is 2×10/(2×10+2×5)=2/3. If, in addition, a dimer with interface #1 is found to be stable, then CSS for interface #1 is max(1,2/3)=1

4.5.1.1 Dimer interface of DpsA and DpsC

The dimer interface is created from two subunits that adjoin along the length of the four-helix bundle in an anti-parallel manner. Comprised of the A and B helices from each subunit, the A helix of one monomer is opposite the B helix of the symmetry related subunit. The BC helices also form part of the dimer interface and are eventually found on the external

surface of the dodecamer (Figure 4.6). The formation of this dimer results in what can be termed an inter-subunit four-helix bundle, creating an environment analogous to that of the intra-subunit ferroxidase centre of ferritin. However, there are only six dimer formations that compose the full dodecamer and these are also vitally important in donating residues involved in creating the ferroxidase centres (FOCs), of which two are found per dimer interface.

DpsA has a dimer interface area of approximately 1310.3 $Å^2$ per monomer, a value that equates to 14.9 % of the total solvent accessible area present on the whole of the monomer. Each monomer has 36 residues involved in this interface region. The solvation free energy gain for the formation of the interface is a negative value of -7.9 kcal/mol; in effect Gibbs free energy is negative, which suggests that the interface is hydrophobic in nature and that the formation of this interface would occur through a spontaneous process. Furthermore, this value does not satisfy the formation of hydrogen bonds across the interface, of which there are 14, and the 8 salt bridges, which carry values of -0.44 kcal/mol and -0.15 kcal/mol respectively per bond formation. All the residues that form salt bridges and hydrogen bonds across the interface are shown in Appendix 14. It can be suggested that each dimer interface has a total Δ^iG of -15.26 kcal/mol, which satisfies hydrogen bonds and salt-bridges. Per dodecamer, this equates to -91.56 kcal/mol.

DpsC has values that are greater than DpsA. A larger interface area of 1870.5 $Å^2$ per monomer has a Δ^iG that corroborates the hydrophobic nature of this interface. At -12.8 kcal/mol, this dimer interface can be considered more hydrophobic than the interface found in DpsA. It also accounts for more of the solvent accessible area present on the DpsC monomer at 19.5 %. Moreover 51 residues of the total protein are involved directly at the interface. This allows for a greater number of hydrophilic interactions: hydrogen bonds and salt bridges. Containing 27 hydrogen bonds and 27 salt bridges, the total Δ^iG that satisfies Hbs and Sbs can be assumed to be -28.73 kcal/mol per dimer and -172.38 kcal/mol for the dodecamer. Consequently, the DpsC dimer interface is stabilised mostly by hydrophilic interactions. Residues involved in hydrogen bond and salt bridge formation in the DpsC dimer interface are shown in Appendix 15. What must also be taken into account is the Δ^iG P-value, which in both cases is greater than 0.5 suggesting the interface is not as hydrophobic as it could be. The data also point toward a role for the N-terminal tail of DpsC. Residues 28 to 33 of the DpsC N-terminal tail are involved in the interface, almost "hugging" the opposite subunit of the dimer. Additionally some of these residues are involved in hydrogen bonding and salt bridge formation. Notably Arg-28 which forms two hydrogen bonds from [NH1] and [O] to the [OE2] and [N] of Glu-135 respectively of its dimer partner. Furthermore, Arg-28 also forms a salt bridge between [NH1] and the [OE2] of the aforementioned Glu-135. This highlights how this region of the tail is "anchored" in place across the surface of the dodecamer, acting as a brace,

and allows the rest of the tail to move freely into solution hence the reason it cannot be resolved in the crystal structure.

4.5.1.2 N-terminal trimer "Ferritin-like" interface of DpsA and DpsC

Unlike the dimer interface that has been described, the N-terminal interface is slightly more complex in terms of the symmetry. Comprised of an arrangement of three symmetry related monomers, each subunit interacts with two symmetry-related neighbouring subunits. This means that there are actually two sets of interfacing residue profiles that can be generated and described as a subunit's interaction with either its clockwise symmetry related partner or its anticlockwise symmetry related partner. Shown in Figure 4.7 are the interfacing regions as a whole highlighted on all three of the subunits. It is easily seen how the interfacing regions, highlighted by the red, green and pink colours, can be divided in two; for example areas where the red is interfacing green residues (from the blue subunit), and areas where it is interfacing with pink residues (from the gold subunit). It is of course the interaction of each subunit with two other subunits that consequently hold together the trimers. This arrangement is also vitally important to the function of the protein; it is this organisation of subunits that creates a channel, akin to those found in ferritin proteins (Bellapadrona et al., 2009), and allows the connection between the solvent and the cavity, thus allowing a route for iron to gain entry to the FOC and the cavity of the protein (Bellapadrona et al., 2010).

The DpsA N-terminal interface (Appendix 16) is comprised of residues predominantly of the CD loop and residues along the length of the D helix. Both tails have sections that contribute to the interface with residues 8-14 of the N-terminal tail and 165-168 of the C-terminal tail. In both cases residues of the tail regions are involved in hydrophilic interactions; residues Lys-8 to Gly-165 at a distance of 2.77 Å are particularly important as this region essentially links the N-terminal tail of one subunit to the C-terminal tail of the other. In addition Thr-10 interacts with Arg-158, linking the N-tail with the D helix with its symmetry related partner and Pro-12 to Arg-126 links the N-terminal tail to the C helix. To paraphrase this, the N-terminal tail of one subunit interacts with the C helix, the D helix and the C-tail all originating from an adjacent subunit. There are a further four hydrogen bonds which strongly link the CD loop and the N-terminal region of the D helix (residues Asp-136 x2, Ala-138 and Asp-141) with a single residue on the C helix (Arg-126 x2) and a single residue on the D helix (Arg-148 x2). In total there are 8 hydrogen bonds and 6 salt bridges which help to stabilise this interface. As this assembly creates a pore, a proposed entry route for iron, the high number of hydrogen bonds and salt bridges reflect the hydrophilic nature of the region.

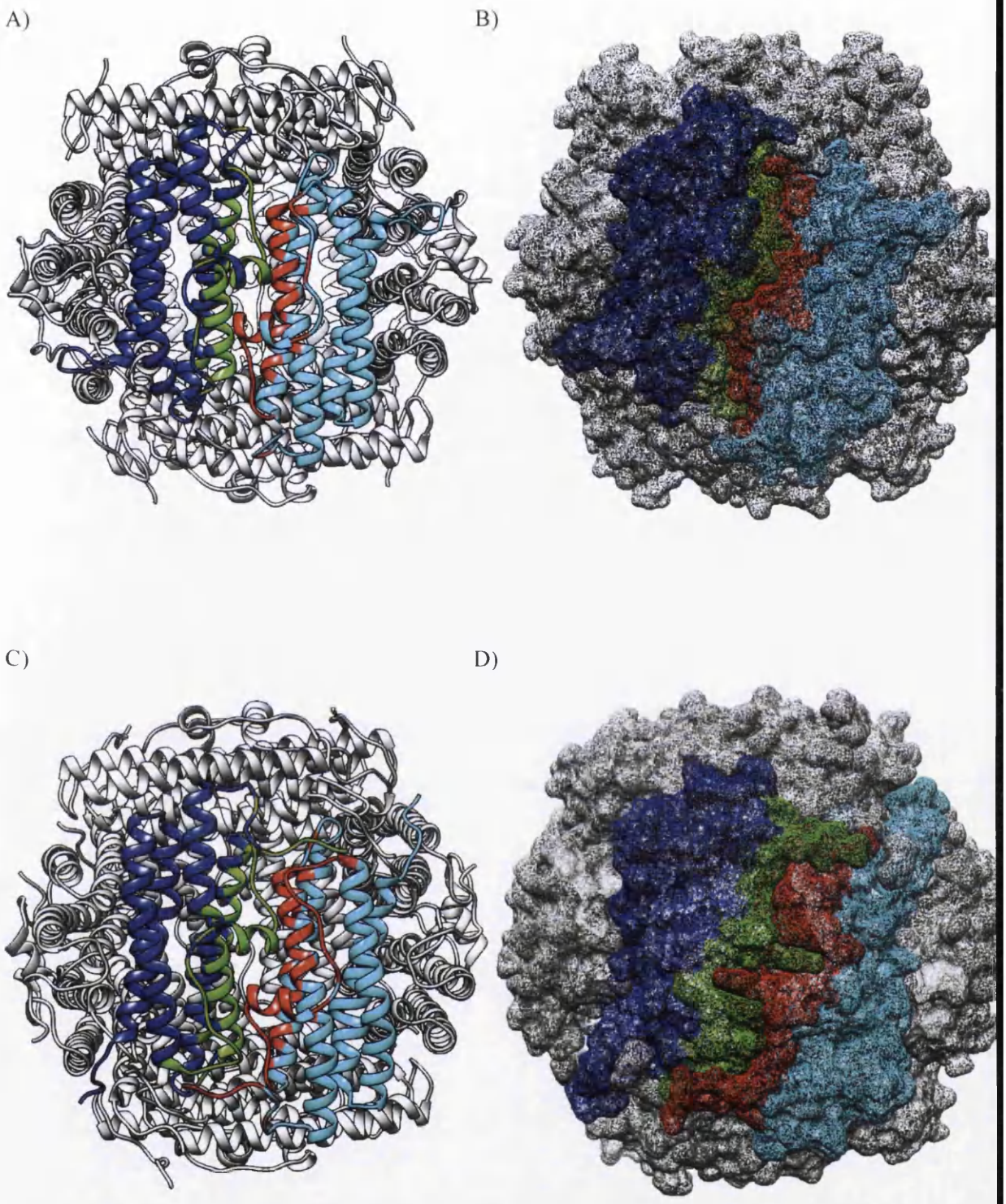


Figure 4.6: Crystal structure models of dodecamers highlighting the dimer interface. Anti-parallel subunits of a single dimer represented in differing colours (blue and turquoise). Highlighted in red and green are the opposing amino acids involved in forming the dimer interface. The other subunits are discoloured for emphasis. DpsA has 36 amino acids per subunit at the interface (A) ribbon (B) mesh surface. DpsC has 51 residues that are interfacing per subunit (C) ribbon (D) mesh surface. Images were created using UCFS Chimera 1.7rc

As a result of the hydrophilic nature of the interface, which actually involves 39 residues per monomer, the interface is the least hydrophobic of DpsA with a Δ^iG of -3.5 kcal/mol. At 742.0 $Å^2$ per monomer, this interface is roughly half the area of the dimer interfaces and follows the trend by having lower complexation significance.

Contrary to DpsA, the N-terminal interface of DpsC is the smallest of its interfaces in terms of area, at 735 $Å^2$ this interface accounts for just 7.9 % of the solvent-accessible area of the protein monomer (Figure 4.7C+D) but similarly to DpsA contains 42 residues. This interface is also predominantly hydrophilic in nature as evidenced by the high number of salt bridges and hydrogen bonds present at the interface: 9 and 8 in quantity respectively. This is in agreement with the function of the region, creating a pore for iron entry, and is reflected by the low value of Δ^iG, - 4.9 kcal/mol. Its high Δ^iG P-value indicates that it is less hydrophobic than it could be, pointing to the evolutionary selective pressure to keep this region hydrophilic in nature. The N and C-terminal tails of DpsC play a slightly diminished role in this interface when compared to DpsA. Only 7 residues of the N-terminal tail are involved in the interface, and from these residues there are two hydrogen bonds and a salt bridge, which help preserve the interface. A hydrogen bond and salt bridge can be formed between Lys-37 of the N-terminal tail and Glu-190 of the D helix of an adjacent subunit. Ala-42 is also capable of forming hydrogen-bonding interactions with Glu-148 of the C helix where the N-terminal tail partially loops over the neighbouring subunit allowing a distance suitable for such hydrophilic interactions. Other hydrogen bonds and salt bridges, listed in Appendix 17, are interactions between residues located on the D helices. Of the total 17 electrostatic interactions formed at the interface, 15 of them result from residues located on the D helix and encompass Asp-167, Asp-171, Glu-176, Arg-179, Glu-182 and Glu-190 highlighting the importance of this helix to the N-terminal "ferritin like" interface.

4.5.1.3 C-terminal trimer interface of DpsA and DpsC

This structure is commonly referred to as the "Dps-like" interface and its assembly, similarly to the N-terminal interface, can potentially create a feature resembling a pore into the cavity of the protein; however divergence in structure and function at this region remains noticeable amongst Dps (Franceschini et al., 2006). For DpsA and DpsC this rule appears to be adhered to, with stark differences shown in the structure and statistics for this region. Figure 4.8 details how the interfacing regions are positioned within the dodecameric assembly of the proteins.

The C-terminal trimer complex of DpsA (Appendix 18) is the smallest interface found within the protein assembly in terms of surface area at 512 $Å^2$ per monomer. The interface is

created with 26 residues from each monomer. These residues have a solvent-accessible area of 516.4 $Å^2$ per monomer; 5.9 % of the total solvent-accessible area of the monomer. With a Δ^iG of -7.1 kcal/mol, it is for the most part hydrophobic in nature and the Δ^iG P-value suggests that the interface is more hydrophobic than if random amino acids were used, thus highlighting how the interface interactions are of a specific nature. However there are still 2 hydrogen bonds (Ala-159 to Gly-49 and Glu-162 to Arg-104) and 2 salt bridges (both Glu-162 to Arg-104), which strengthen the interface. These residues are located at the C-terminal region of the D helix, the C-terminal tail of one subunit and the loop between the BC helix and the C helix of the adjacent subunit and as such are located on the surface of the dodecamer. This allows at least part of the C-terminal tail to have a role in the oligomeric assembly. However the latter region of the tail from residue 163 onwards actually bends away from the adjacent subunit and back towards its own four-helix bundle.

The C-terminal trimer interface of DpsC is predicted to play a more significant role in maintaining oligomeric assembly than that of the one found in DpsA. In DpsC, all the four major helices lend residues towards the interface; residues 73-75 of the A helix, residues 81-88 and 90 of the B helix with the C helix lending residues 136-141 while the D helix lends residues 186 and 187 along with 190-192. In addition to these residues, the entire loop between the A and B helices is interfaced; furthermore, the entire DpsC C-terminal tail is involved in the interface bridging over the adjacent subunit as seen in Figure 4.8C where the green residues of the blue subunit overlap the red residues of the turquoise subunit creating an interfacing region. The interface has an area of 820.4 $Å^2$ per monomer and is created with 33 residues in total per subunit. The solvent-accessible area of the interface is 9.0 % of the total solvent-accessible area of the monomer. Whilst having an extremely negative Δ^iG of -14.2 kcal/mol, this highly hydrophobic interface is also strengthened by 5 hydrophilic interactions involving hydrogen bonds (Appendix 19). Of significant interest is the hydrogen bond between His-199 [NE2] and Pro137 [O] of an adjacent subunit. This anchors the tail to the scaffold of the dodecamer allowing it to support and stabilise the assembly of the dodecamer. Since this is a three-fold symmetry related structure, the tails triangulate the entire interfacing region adding weight to their function as ligatures of the dodecameric assembly. Additionally, the CSS score for this interface is also the highest of all the CSS scores, thus it is significantly important to the assembly of the dodecamer. This comparison between the C-terminal interfaces of DpsA and DpsC can be seen in Figure 4.8.

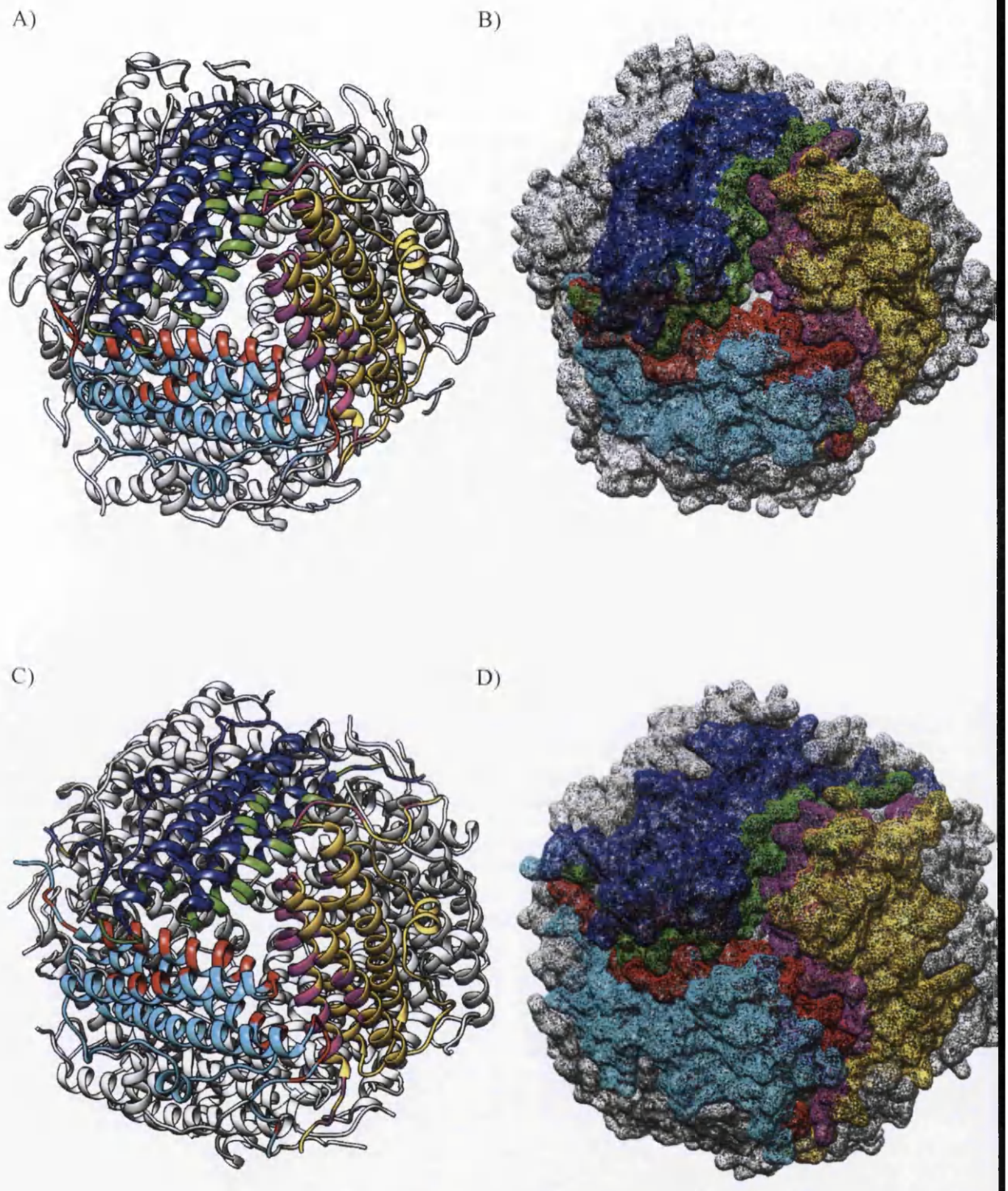


Figure 4.7: Crystal structure models of dodecamers highlighting the N-terminal trimer interface. This structure forms the "Ferritin-like" pore. 3-fold symmetry related monomers are represented by the colours blue, turquoise and gold, with the interfacing residues of the subunits coloured in green, red and pink. The other subunits are discoloured for emphasis. (A) DpsA ribbon, (B) DpsA mesh surface, (C) DpsC ribbon and (D) DpsC mesh surface. Images were created using UCFS Chimera 1.7rc

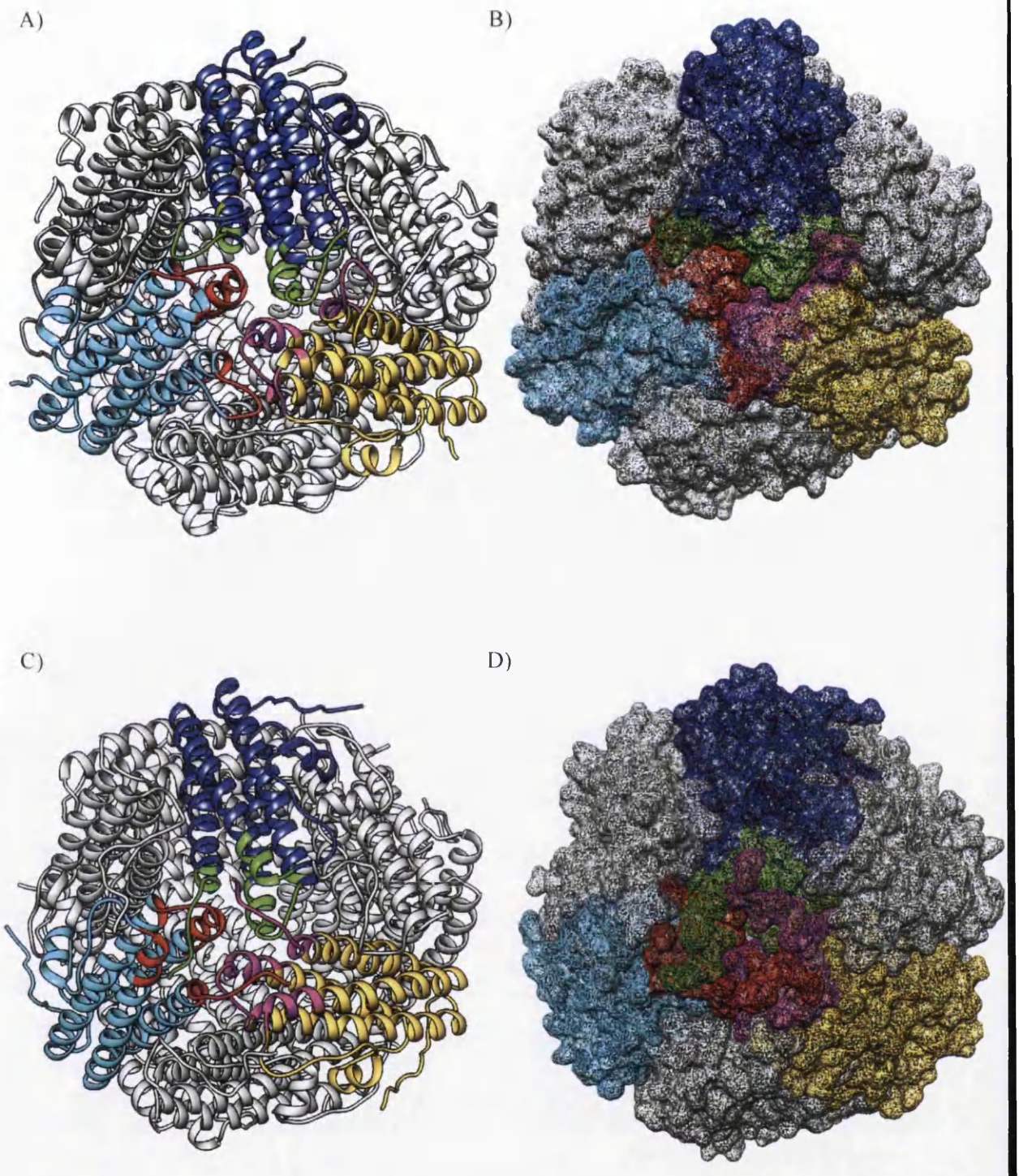


Figure 4.8: Crystal structure models of dodecamers highlighting the C-terminal trimer interface. This structure forms the "Dps-like" pore. 3-fold symmetry related monomers are represented by the colours blue, turquoise and gold, with the interfacing residues of the subunits coloured in green, red and pink. The other subunits are discoloured for emphasis. (A) DpsA ribbon, (B) DpsA mesh surface, (C) DpsC ribbon and (D) DpsC mesh surface. Images were created using UCFS Chimera 1.7rc.

4.5.2 PISA comparisons

The statistics from the PISA analysis of DpsA and DpsC can be used to evaluate how these proteins compare to other Dps proteins with structures within the PDB. A variety of Dps were used for analysis including the well-studied *E. coli*, *M. smegmatis* and *L. innocua* Dps. The PDB IDs were individually submitted to the PISA server, all ligands de-selected and the interfaces verified manually before downloading and compiling the statistics. The table is provided as Appendix 20. Intriguingly, the PISA results allow this set of proteins to be separated into two subgroups; those with high numbers of hydrogen bonds and salt-bridges across the dimer interfaces and those without such high numbers. Furthermore, these groups also reflect the classification of the organisms as either a mesophile or extremophile whereby any Dps with high numbers of Hb or Sb across the dimer interface originate from extreme environments. There is also a tendency for this interface to be extended in surface area. As would be assumed, DpsA from mesophilic *S. coelicolor* falls into the mesophilic group based on its average number of Hb and Sb across the dimer interface. However, DpsC proves to be more enigmatic. Containing an average of 27 Hb and 27 Sb across the dimer interface, and with an extended surface area of 1870.5 $Å^2$, this protein would be grouped with the extremophiles such as the halophilic archaeon *H. salinarum*. This demonstrates the importance the dimer interface may have to the overall stability of the assemblies.

4.5.3 Role of Dps tails maintaining assemblies

Taking all the PISA analyses into consideration it is possible to collate the data in order to exhibit the contribution that the N- and C-terminal tails give to maintaining the interface assemblies. Taking each tail at a time, the interface they are involved with and specifically the amino acids with which they interact can be analysed. This allows a full understanding of their contribution to oligomeric assembly to be gained.

4.5.3.1 DpsA N-terminal tail

There are 12 amino acids from the N-terminal tail that have been resolved within the crystal structure. This tail has been found to be involved in two of the three structural elements defined by the PISA analysis; the dimer interface and the N-terminal trimer interface. Table 4.5 highlights the interface in which each residue contributes to and the exact amino acids with which they interface from a single symmetry related subunit. Also noted are the amino acids with which they form predicted hydrogen bonding interactions (Hb). For the case of the dimer interface, there is only a single symmetry related partner, and for the N-terminal trimer interface, there are two partners. However the N-tail only interacts with a single symmetry related subunit, in a clockwise manner around the pore. Using the illustrative colours in Figure

4.9A, it can be seen how the blue subunit's N-tail interacts with the C-tail from the turquoise subunit. In turn the turquoise subunit's N-tail interacts with the C-tail of the gold subunit, and the circle is complete when the gold subunit's N-tail interacts with the blue subunits C-tail.

Table 4.5: DpsA N-tail residue inter-subunit interactions

Given below are the N-tail residues and the residues that they interact with from a single symmetry related subunit per interface.

N-tail Residue	Dimer Interface	N-terminal trimer	Total Interfacing Residues
Asp-4	-	-	**0**
Leu-5	-	-	**0**
Thr-6	-	-	**0**
Pro-7	Arg-104	-	**1**
Lys-8	-	Gly-165 (Hb), Gly-166, Ala-167	**3**
Tyr-9	-	Arg-158, Glu-162, Gly-166, Leu-168	**4**
Thr-10	-	Arg-158 (Hb)	**1**
Val-11	-	Glu-151	**1**
Pro-12	-	Thr-119, Val-122, Glu-123, Arg-126 (Hb), Gln-154,	**5**
Gln-13	-	Arg-126	**1**
Ile-14	-	Arg-126	**1**
Glu-15	-	-	**0**
		Total Interfacing residues with N-tail	**17**

There are a number of observations that can be made from the table with use of the molecular modelled interfaces. Lys-8 and Tyr-9 are both heavily involved in interacting with the C-terminal tail of an adjacent subunit within the N-terminal trimer interface. Lys-8 can be found interfacing with three C-tail residues (Gly-165, Gly-166 and Ala-167) and hydrogen bonds with Gly-165 (Figure 4.9A). Tyr-9 is interfaced with residues Arg-158, Glu-162, Gly-166 and Leu-168, with all but Arg-158 constituting the C-tail. Pro-12 interfaces with 5 different residues (Thr-119, Val-122, Glu-123, Arg-126 and Gln-154) also accepting hydrogen bonding from Arg-126. Pro-12 is a hydrophobic residue and thus its location on the surface of the protein would be stabilised by the shielding achieved by the five residues with which it interfaces (Figure 4.9B).

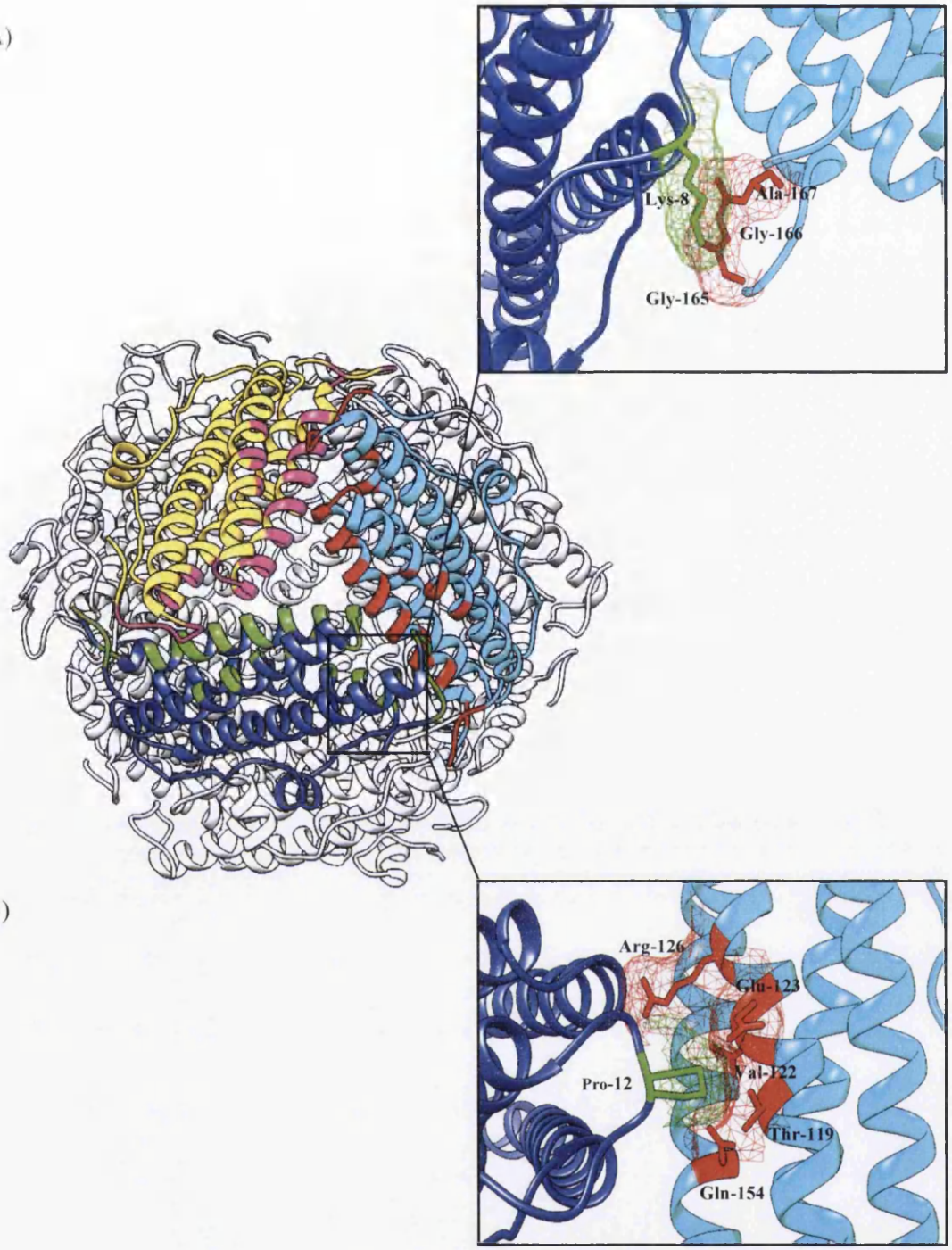


Figure 4.9: Important residues at the DpsA N-terminal trimer interface. (A) Interfacing residues of DpsA N-tail Lys-8 (green) interfaces with three residues (Gly-165, Gly-166 and Ala-167 – all red) from the C-terminal tail of a symmetry related subunit at the N-terminal trimer interface. Lys-8 also hydrogen bonds with Gly-165.

(B) Interfacing residues of DpsA N-tail Pro-12 (green) is interfacing with five residues (Thr-119, Val-122, Glu-123, Arg-126 and Gln-154 – all red) from an adjacent symmetry related subunit of the N-terminal trimer interface. Hydrogen bonding interactions also form between Pro-12 and Arg-126. Images were created using UCFS Chimera 1.7rc

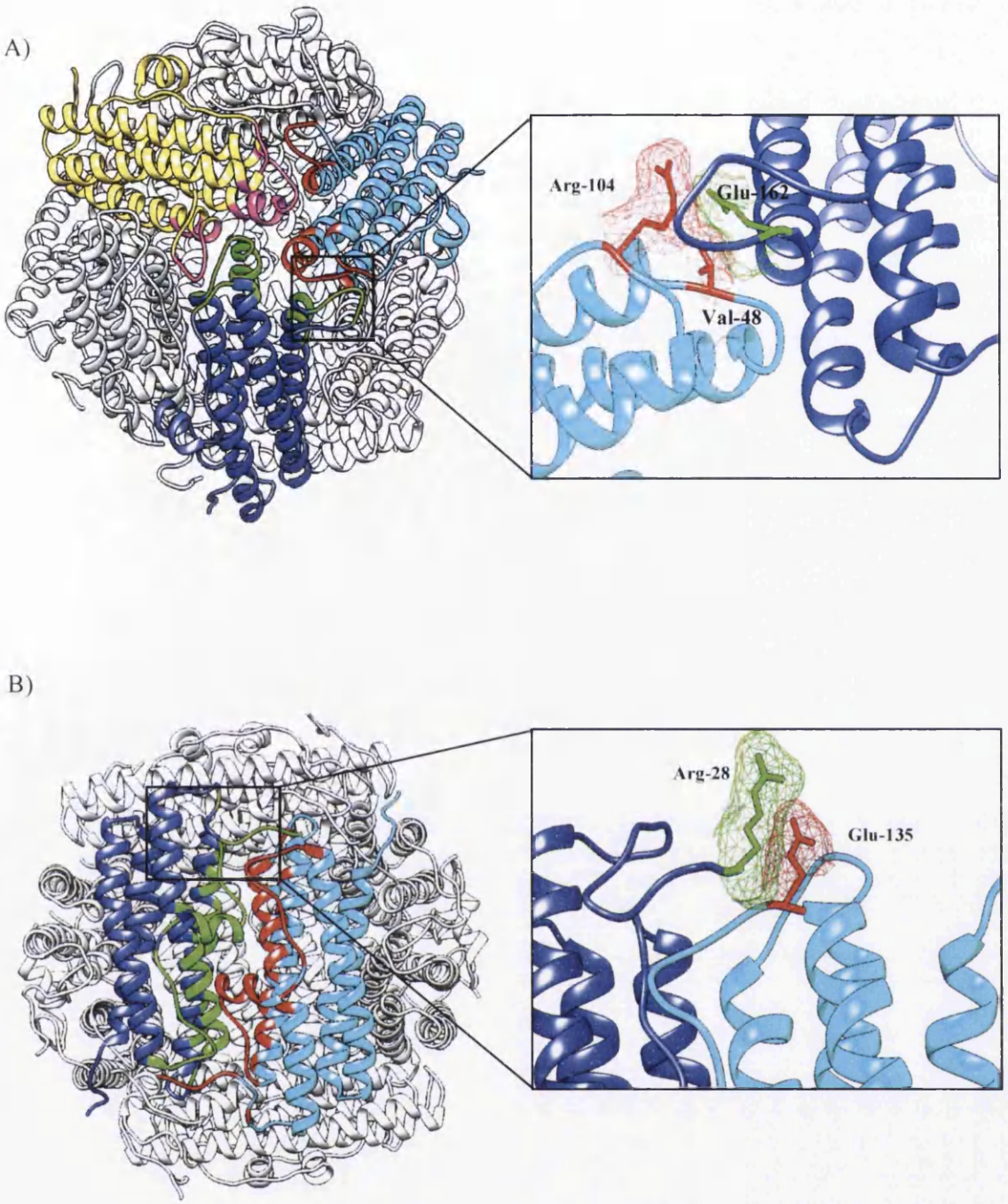


Figure 4.10: Important residues at DpsA and DpsC interfaces. (A) Interfacing residue of DpsA C-tail Glu-162 (green) interfacing with Val-48 and Arg-104 (red) at the C-terminal trimer interface. Hydrophilic interactions occur between Glu-162 and Arg-104 to strengthen the interface.

(B) Antiparallel dimer of DpsC. Residue Arg-28 (green) of the N-tail interfaces with Glu-135 (red) of the adjacent monomer creating an environment with distances conducive to the formation of hydrophilic interactions.

4.5.3.2 DpsA C-terminal tail

Within the crystal structure of DpsA, 9 residues of the C-terminal tail have been resolved. The structure and PISA analysis demonstrate how this tail is involved with two interfacing assemblies: the N-terminal trimer and the C-terminal trimer. The residues interfacing the C-terminal tail within the N-terminal trimer interface are only from the N-terminal tail. Within the C-terminal trimer interface, five C-tail residues all interface with Arg-104, which is situated in the loop between the BC helix and the C-helix. This suggests that this residue is important in maintaining the C-terminal tail within position. The interface of Glu-162 at the C-terminal trimer interface is shown in Figure 4.10A.

Table 4.6: DpsA C-tail residue inter-subunit interactions

C-Tail Residue	N-terminal trimer	C-terminal trimer	Total Interfacing Residues
Glu-162	Tyr-9	Val-48 Arg-104 (Hb, Sb x 2)	**4**
Ser-163	-	Val-48, Arg-104	**2**
Ala-164	-	Arg-104, Ala-105	**2**
Gly-165	Lys-8 (Hb)	Arg-104	**2**
Gly-166	Lys-8, Tyr-9	Arg-104	**2**
Ala-167	Lys-8	-	**1**
Leu-168	Tyr-9	-	**1**
Ala-169	-	-	**0**
Ala-170	-	-	**0**
		Total Interfacing Residues with C-tail	**14**

4.5.3.3 DpsC N-terminal tail

The N-tail of DpsC, which has residues 28 to 44 resolved within the crystal structure, contributes to the formation and stabilisation of both the N-terminal trimer interface and also the dimer interface. The residues and any hydrophilic interactions are presented in table 4.7. A substantial 29 residues are interfaced with the N-tail highlighting its importance to the assembly and stabilisation of DpsC.

The N-tail of each dimer enfolds its partner; acting as a bracket to strengthen the interface. Five residues are found to interface with the opposing subunit with a total of five hydrogen bonding interactions and a salt bridge predicted to form across the interface from the residues of the N-tail. Of particular importance would appear to be Arg-28, the first residue of DpsC resolved in the structure visible likely because of the hydrophilic interactions with Glu-135 (Figure 4.10B). Val-133 would also appear to be a key residue in this interface as it is

involved with three of the five N-tail residues interfacing. Since Val-133 is a hydrophobic amino acid, the residues of the N-tail help to shield it from the external environment and this could be fundamental in creating the tail topology helping to draw the N-tail across the surface of the monomer in which Val-133 is located.

Seven residues of the N-tail play a role in the N-terminal trimer interface with a total of four electrostatic interactions strengthening the interface. These residues are generally different to the residues involved with the dimer interface. However Ile-30 is involved in both interfaces with interfacing with a total of five amino acids. Furthermore as another hydrophobic residue present on the surface of the protein, it would require shielding. This could in fact provide a method for correctly positioning the tail in place on the surface of the dodecamer allowing other vital interactions to occur.

Table 4.7: DpsC N-tail residue inter-subunit interactions

N-tail Residue	Dimer Interface	N-terminal trimer	Total Interfacing Residues
Arg-28	Glu-135 (2Hb) (Sb)	-	**1**
Thr-29	Glu-134, Val-133, Gly-132	-	**3**
Ile-30	Glu-135, Arg-77, Val-133 (Hb)	Thr-195, Pro-196	**5**
Gln-31	Gly-132, Arg-77, Val-133 (2Hb)	-	**3**
Glu-32	-	-	**0**
Phe-33	-	-	**0**
Gly-34	-	-	**0**
Thr-35	-	-	**0**
Val-36	-	-	**0**
Lys-37	Arg-77	Glu-190 (Hb), (2Sb)	**2**
Gln-38	-	-	**0**
Phe-39	-	Trp-186, Ala-189, Glu-190	**3**
Pro-40	-	Leu-144, Glu-148, Ala-185	**3**
Val-41	-	Glu-148, Leu-151, Glu-182	**3**
Vla-42	-	Glu-148 (Hb), Leu-151, Thr-152, His-155	**4**
Leu-43	-	His-155, Glu-182	**2**
Thr-44	-	-	**0**
		Total Interfacing Residues with N-tail	**29**

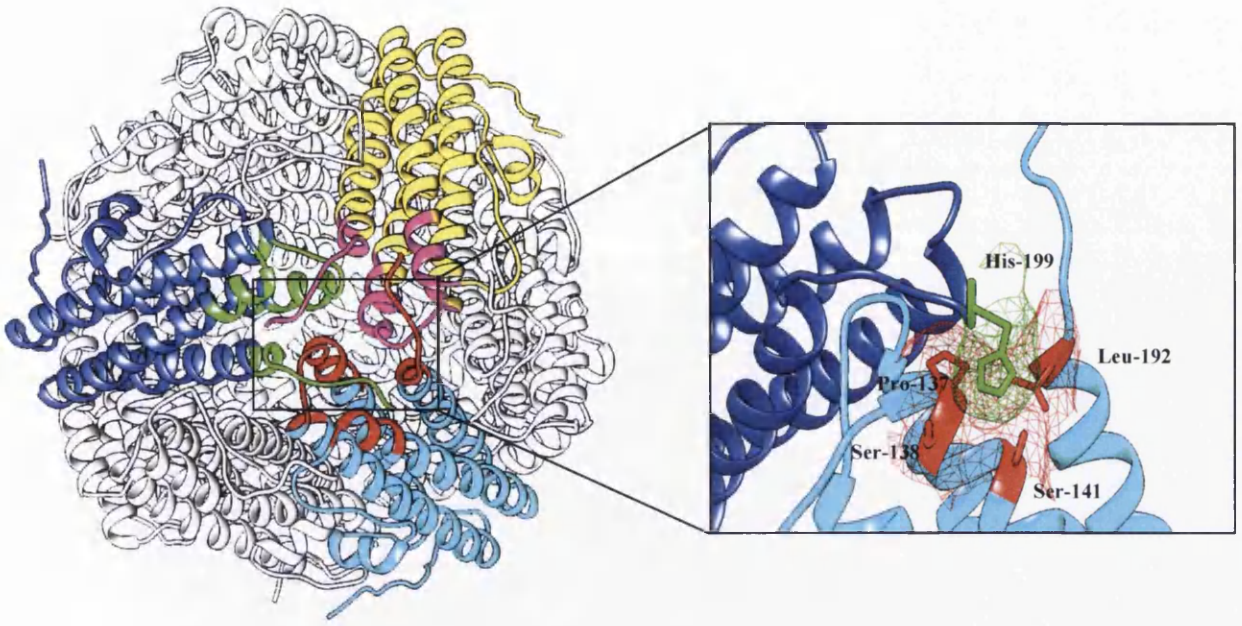


Figure 4.11: DpsC C-terminal tail interface at the C-terminal trimer interface. His-199 (green) of the C-terminal tail of one subunit interfaces with Pro-137, Ser-138, Ser-141 and Leu-192 (red) of an adjacent three fold symmetry related subunit, hydrogen bonding interactions with His-199 help to strengthen this interface.

4.5.3.4 DpsC C-terminal tail

The C-tail of DpsC is a lot shorter than the N-tail. With the seven residues resolved within the structure there is a total number of 18 interfacing residues split between the N-trimer interface and the C-trimer interface. The residues of the C-tail involved in formation of the N-trimer interface are only mildly interfacing and are at the boundary of distances that qualify as an interface from the PISA analysis and as such would not contribute substantially to the stability of the interface.

The impact the C-tail has on stabilising the C-terminal trimer is of greater significance. The three fold symmetry displayed at the C-terminal pore allowing three subunits to interact lets the C-tail of each subunit interact with a single other subunit, in an anti-clockwise circular manner. Perhaps the most important is the interaction of His-199, which is predicted to hydrogen bond with Ser-138, pinning the tail to the surface of the threefold symmetry related subunit.

Table 4.8: DpsC C-tail residue inter-subunit interactions

C-Tail Residue	N-terminal trimer	C-terminal trimer	Total Interfacing Residues
Val-193	Ile-30, Lys-37	Gly-78, Arg-77	**4**
Asp-194	-	-	**0**
Thr-195	Ile-30, Lys-37	Gly-78, Arg-77, Val-136	**5**
Pro-196	Ile-30, Lys-37	Glu-135, Val-136, Pro-137	**5**
Leu-197	-	Pro-137	**1**
Val-198	-	Pro-137, Leu-192, Asp-194	**3**
His-199	-	Ser-138, Pro-137, Ser-141 (Hb) Leu-192	**4**
		Total Interfacing Residues with C-tail	**22**

4.5.4 Prediction of Interface Hotspots using Computational Alanine Scanning

Computational alanine scanning was carried out to identify residues within the protein-protein interfaces that are energetically important. From here on, these residues will be referred to as "hot spots". A freely available online server, (available at http://robetta.bakerlab.org/alaninescan) requires a multi-chain protein complex PDB file and the interfacing protein chains to be defined before computation can be carried out. The alanine scanning process represents a computational approach of "alanine scanning mutagenesis" where each residue of an interface is individually mutated to alanine, and the effect this has on the affinity of a protein-protein interface measured (Kortemme and Baker, 2002). The main result generated by the *in silico* analysis is a predicted change in binding free energy upon alanine substitution (Kortemme et al., 2004). Residues that return a $\Delta\Delta G$ (complex) which have a predicted binding free energy change value equal to or greater than 1 kcal/mol are termed "hot spots". The residues involved in the interfaces to be tested have already been established by the PISA analysis method, however, to further certify the PISA analysis, the alanine scanning process also predicts residues involved in the interface between the pre-selected structural chains. Furthermore it can also reveal if any hot spots are predicted in the tail regions. The only other Dps to be investigated with alanine scanning is the *E. coli* Dps (*Ec*Dps) (Zhang et al., 2011). There are a number of differences found between this protein and the *Sc*Dps. The most obvious difference is an absence of aromatic residues (Phe, Tyr and Trp) at the Dps-like pore of *Ec*Dps. Both DpsA and DpsC have multiple aromatic residues involved with this interface; furthermore these residues were predicted to be hot spots. Moreover the C-terminal interfaces of both DpsA and DpsC are richer in hot spot residues when compared to *Ec*Dps.

4.5.4.1 DpsA Alanine Scanning Hot spots

The results for DpsA are split into each interface starting with the dimer interface, followed by the N-terminal trimer and the C-terminal trimer

Table 4.9 Hot spot residues predicted by alanine scanning of DpsA dimer interface

Residue	ΔΔG (complex) kcal/mol	Location on monomer
Leu-38	1.39	A – helix
His-42	1.52	A – helix
Trp-45*	3.02	A – helix
Gln-85	1.26	B to BC loop
Thr-87*	1.87	B to BC loop

*Residues labelled with * are predicted PISA analysis to form hydrophilic interactions by.*

Table 4.10 Hot spot residues predicted by alanine scanning of DpsA N-terminal trimer interface

Residue	ΔΔG (complex) kcal/mol	Location on monomer
Lys-8*	1.31	N – tail
Tyr-9	2.67	N – tail
Arg-76*	4.89	B – helix
Arg-126	2.47	C – helix
Thr-139	3.24	D – helix
Asp-141*	1.58	D – helix
Arg-148*	3.14	D – helix
Glu-151*	1.03	D – helix
Gln-152	2.48	D – helix
Gln-154	2.15	D – helix
Trp-155	2.38	D – helix
Arg-158	3.20	D – helix

*Residues labelled with * are predicted by PISA analysis to form hydrophilic interactions.*

Table 4.11 Hot spot residues predicted by alanine scanning of DpsA C-terminal trimer interface

Residue	ΔΔG (complex) kcal/mol	Location on monomer
Val-48	1.08	A to B loop
Phe-52	2.65	B – helix
Ile-53	1.59	B – helix
Arg-104*	1.52	BC to C loop
Trp-155	1.32	D – helix
Phe-156	1.36	D – helix
His-160	1.40	D – helix

*Residues labelled with * are predicted by PISA analysis to form hydrophilic interactions.*

There are a total of 23 predicted hot spots which are suggested to be energetically important to the maintenance of the protein interfaces and thus the oligomeric assembly overall. Five at the dimer interface, twelve at the N-terminal trimer interface and seven at the C-terminal trimer interface. Out of these 23 residues, 8 are involved in interactions across the interfaces (hydrogen bonding interactions or salt bridges). The hot spot residue Trp-155 is the only residue that has a ΔΔG (complex) value greater than 1.0 kcal/mol that is identified at two types of interface; the N-terminal trimer and the C-terminal trimer interfaces. This residue would be an ideal candidate for disruptional mutagenesis studies. The D helix would appear to be important to oligomeric assembly supplying 10 different residues that are regarded as hot spots. There are only two tail residues predicted as hot spots; Lys-8 and Tyr-9. Both these residues are present on the N-tail with Lys-8 implicated in hydrogen bond formation previously described (Figure 4.9A). This substantiates the importance of the N-tail to oligomeric assembly of DpsA and corroborates the results pertaining to the *in vitro* assembly of DpsA deleted tail variant proteins, which are described in Chapter 5.

Whilst alanine scanning does not take account of the symmetry of a protein complex, it is assumed that the binding free energy values for each mutation are additive. We should therefore consider the ΔΔG (complex) value for Arg-76. Involved in a three-fold symmetry related interface, Arg-76 has the highest ΔΔG (complex) value for DpsA (4.89 kcal/mol). There are three Arg-76 residues involved per N-terminal trimer interface, and there are four N-terminal trimers per dodecamer, thus a mutation of Arg-76 would result in a ΔΔG (complex) value of 8.68 kcal/mol per dodecamer. Because of this, Arg-76 would also be considered an interesting target for *in vitro* mutation and analysis, as its significance is greater than other residues.

4.5.4.2 DpsC Alanine Scanning Hot spots

Table 4.12 Hot spot residues predicted by alanine scanning of DpsC dimer interface

Residue	ΔΔG (complex) kcal/mol	Monomer location
Ile-30*	1.66	N – tail
Arg-63*	1.17	A – helix
Lys-70*	3.63	A – helix
Trp-74*	3.78	A – helix
Leu-75	1.2	A – helix
Asp-100*	2.57	B – helix
Asp-116	1.98	B to BC loop
Arg-118*	4.82	BC – helix
Arg-128*	2.34	BC to C loop
Asp-131*	1.05	BC to C loop
Glu-134	2.81	BC to C loop

*Residues labelled with * are predicted by PISA analysis to form hydrophilic interactions*

Table 4.13 Hot spot residues predicted by alanine scanning of DpsC N-terminal trimer interface

Residue	ΔΔG (complex) kcal/mol	Monomer location
Phe-39	1.66	N – tail
Arg-105*	3.26	B – helix
Glu-148	1.29	C – helix
His-155*	2.74	C – helix
Asp-167*	1.25	D – helix
Asp-171*	1.39	D – helix
Arg-179*	3.39	D – helix
Glu-182*	3.06	D – helix
Trp-186	2.74	D – helix

*Residues labelled with * are predicted by PISA analysis to form hydrophilic interactions*

Table 4.14 Hot spot residues predicted by alanine scanning of DpsC C-terminal trimer interface

Residue	ΔΔG (complex) kcal/mol	Monomer location
Phe-81	2.95	B – helix
Tyr-82*	2.93	B – helix
Gln-83*	3.14	B – helix
Trp-186	1.61	D – helix
Phe-187	1.64	D – helix
His-191*	2.19	D – helix
Leu-197	1.49	C – terminal tail
Val-198	1.24	C – terminal tail
His-199*	2.06	C – terminal tail

*Residues labelled with * are predicted by PISA analysis to form hydrophilic interactions*

There are a total of 29 hot spots predicted within the interfaces of DpsC, with 18 of these residues involved in hydrophilic interactions. Whilst DpsC contains only 6 more hot spots than DpsA, DpsC has more than double the number of hot spot residues involved in hydrophilic interactions highlighting the importance of these interactions to the stability of DpsC. If they were to be mutated they would not only weaken stability but could also disrupt assembly.

DpsC also has the presence of a hot spot on the BC helix (Arg-118). Moreover, this residue also holds the highest ΔΔG (complex) value for DpsC (4.82 kcal/mol). Given that DpsA displayed a disrupted assembly upon deletion of the BC helix (Chapter 5), a point mutation at Arg-118 provides an opportunity to analyse the contribution of the BC helix to assembly of DpsC without deleting the whole helix. Unlike DpsA, both the N and C terminal tails of DpsC contain hot spots. The two hot spots of DpsA are located on the N-tail and thus these are only implicated in the stability of the N-terminal trimer, however DpsC has tail hot spots in each of the three interface types. Confirming the predicted importance of His-199 (Figure 4.11) in bracing the C-terminal trimer interface, this residue returns the highest ΔΔG (complex) of the DpsC tail hotspots. Given its anchoring role, if this residue were to be mutated, it could disrupt the function of the other tail residues particularly the hot spots Leu-197 and Val-198.

As previously stated, the Dps of *E. coli* is the only other Dps that has been investigated by alanine scanning to predict residues, which was then followed up by *in vitro* structural assembly and stability experiments. The focus was placed on the dimer interface and the N-terminal interface since *Ec*Dps displayed no obvious hotspots at the C-terminal interface. A number of hot spot residues, which were experimentally mutated, displayed cage assembly but

presented lower temperature stability. However two residues, which were predicted to be the most energetically important, were highly conserved Arg-83 and Arg-133. Mutations at these positions to alanine resulted in a substantial amount of dimer formation (39% and 49%, respectively) while double mutation resulted in complete dimer formation and no dodecameric assembly. Alignment of a number of Dps protein sequences, including the well-characterised Dps of *E. coli*, *M. smegmatis* and *L. innocua*, highlighted the conserved nature of these residues (alignment given as Appendix 21). In DpsA these residues are Arg-76 and Arg-126. As previously described, Arg-76 is also the highest scoring hot spot for DpsA but there are other residues that are predicted to be more energetically significant than Arg-126. DpsC retains the arginine on the B helix and is numbered Arg-105, however the second arginine is replaced by a histidine (His-155) neither of which are predicted to be the most energetically significant residues by alanine scanning. Furthermore the alignment, which contains predominantly Dps from mesophilic organisms, highlights that the histidine is conserved in the *T. elongatus* Dps that is extreme in origin.

4.5.5 Electrostatic surface potential molecular simulation

The electrostatic surface potential (ESSP) of a protein can facilitate the determination of its interaction with a charged ligand. In the case of the Dps proteins, an interaction with positively charged iron renders the reconstruction of the electrostatic surface potential of the protein an important asset that allows for the visualisation and evaluation of the effect this potential has on the function of the protein. The three structures of importance to the assembly and function of the protein described in the PISA analysis can again be studied from the perspective of the electrostatic forces they employ which gives the protein its functional activity of drawing iron in to the protein cavity via the pores, culminating in the oxidation and deposition of iron as an iron oxide core.

UCSF Chimera was used for the calculation of an electrostatic potential map that is then used to colour the surface of the protein according to electrostatic potential. Colours ranging from red, indicating negative electrostatic potential, to white, which signifies a neutral charge, and on to blue, describing positive electrostatic potential, can then be applied to the surface. The colour-associated values were set to a linear range to allow the minimum and maximum values to be included and a surface offset distance of 1.4 Å was used. The structures were first prepared for electrostatic calculation by adding hydrogens and assigning atomic charges and radii to the relevant atoms. This process was carried out using a PDB2PQR tool (Dolinsky et al., 2007), creating a PQR file; an adapted PDB file with partial atomic charge (Q) and radius (R). Further calculations using the Adaptive Poisson-Boltzmann Solver (APBS) tool were used in order to create an electrostatic potential map (DX file), which was opened as a new model

within the program along with the electrostatic colouring tool. It was found that when the dodecamer was treated as a whole, in effect using the surface of the dodecamer to compute the electrostatic surface potential, the viewing of the cavity was not possible. In order to mitigate this, subunits were independently submitted for calculations and reassembled upon completion as separate models. This ensures that each monomer has its surface fully mapped and allows removal of subunits to reveal the internal surfaces of the cavity.

4.5.5.1 The external surface ESSP

Figure 4.12 displays the dodecamers of DpsA and DpsC aligned at the N-terminal pore and with surfaces coloured according to the electrostatic potential. Obvious differences can be seen in the abundance of positive and negative electrostatic potential signified by the blue and red colours respectively. DpsA has a surface with a slightly greater abundance of positive potential. However, at the ferritin like pore, DpsA has a greater red colour, indicating an area with greater negative potential compared with DpsC. This is in no small part due to the size of the pore. Pro-137 of DpsA is located on the surface of the pore and has a distance of 12.6 Å between the Cα from another three fold-symmetry related Pro-137. This pore size is well within the confines of pore sizes reported for other Dps proteins (Haikarainen and Papageorgiou, 2010). Lys-134 provides a positive charge at the opening of the pore, analogous to other Dps such as *D. radiodurans* (Zeth et al., 2004). This is followed by a series of acidic residues (Glu-130, Glu-140 and Asp-141), which facilitate the negative potential gradient. DpsC has a pore size of 13.7 Å between the Cα of two symmetry related Asp-167 residues, thus opening up the pore and resulting in a spread of the potential created by acid residues lining the pore (Glu-163, Asp-167, Asn-170 and Asp-171). Whilst there is a significant difference in distances at the surface of the pores, the difference at the cavity side of the pore is less defined. At 8.6 Å between the Cα of two Asp-171 residues of DpsC, it is only slightly larger than the 8.5 Å distance between Asp-141 of DpsA, in both cases creating a funnel shape channel.

The C-terminal "Dps-like" pore of both DpsA and DpsC proteins (Figure 4.13) is not only smaller than the N-terminal pore, but is also neutral. The DpsA "Dps-like" pore has distances of 9.99 Å (Pro-50) at the surface and 6.07 Å (Ala-54) at the cavity. These distances are comparable with DpsC which measures 9.53 Å (Ala-79) at the surface and 6.17 Å (Gln-83) at the cavity. Nevertheless what must be taken into account is that the pore residues are measured using the Cα atoms and thus, depending on the frozen rotational conformations adopted by the pore lining residues within the crystals, the dimensions of the channels may differ when conformational changes occur in solution. Whilst the rotational conformations of the His-51 residues of DpsA which protrude into the pore may change. resulting in an "opening up" of the pore, the mildly basic imidazole ring of histidine can facilitate the neutralisation of

electrostatic potential and as a result can still block the route of positive iron from entering the cavity, thus adding weight to the argument that this pore is not ordinarily used for iron translocation in these Dps proteins.

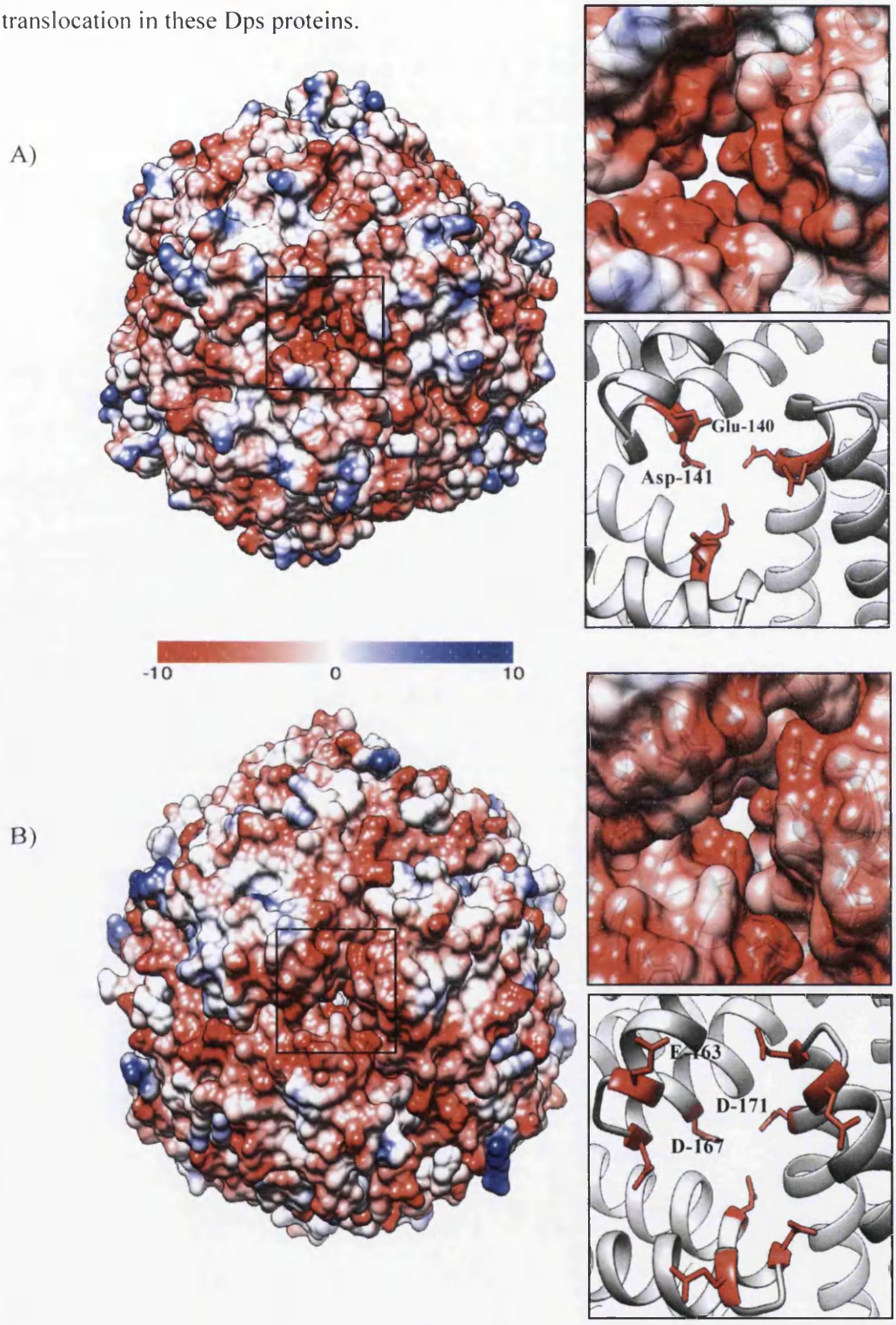


Figure 4.12: Electrostatic surface potential of DpsA and DpsC "ferritin-like" pore. Surface of DpsA (A) and DpsC (B) coloured according to the electrostatic surface potential. Red colour is negative; white is neutral and blue positive. Residues highlighted in DpsA at the "ferritin like pore" are Glu-140 (E and Asp-141, for DpsC Glu-163, Asp-167 and Asp-171 are highlighted. Images were created using UCFS Chimera 1.7rc

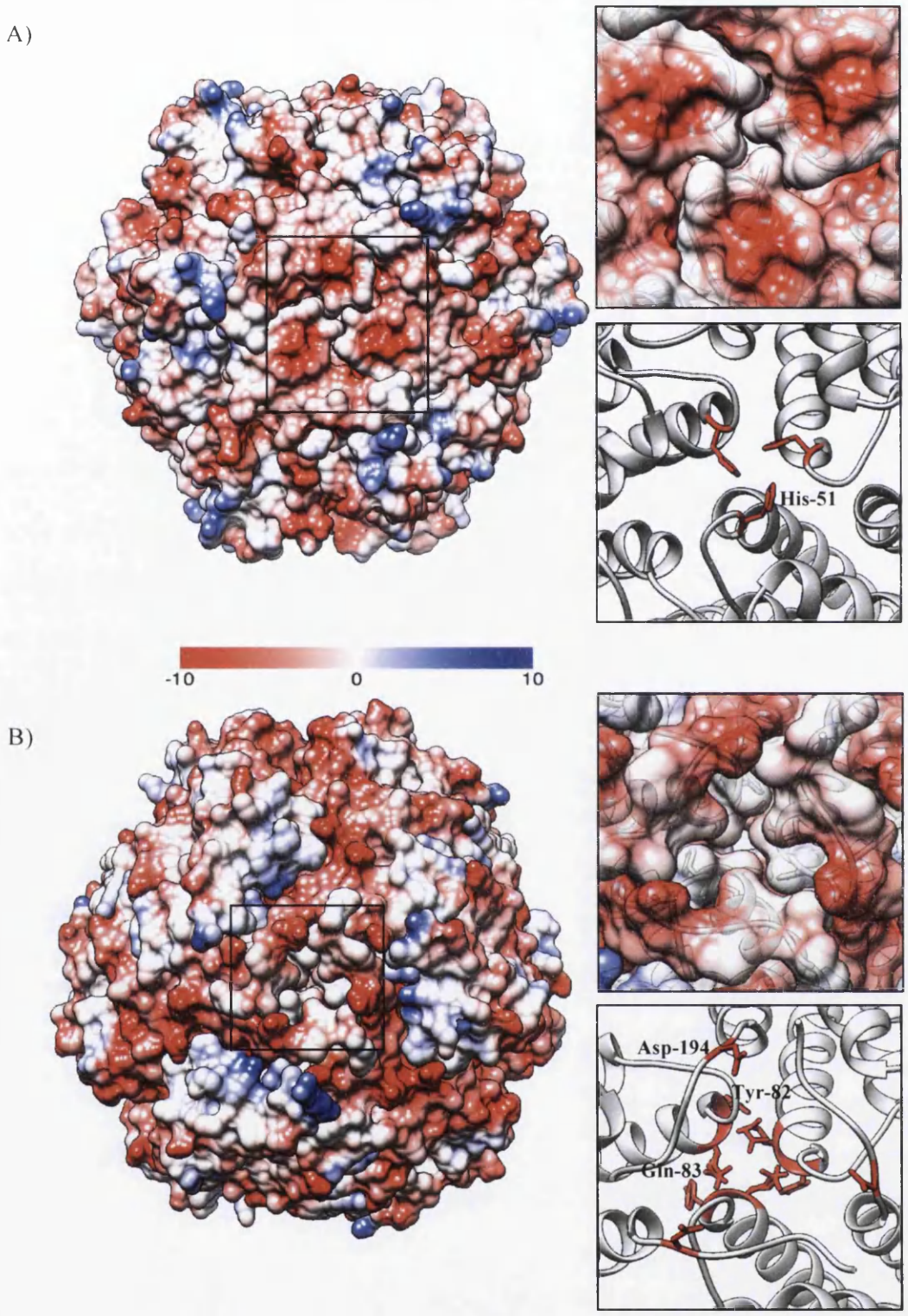


Figure 4.13: Electrostatic surface potential of DpsA and DpsC "Dps-like" pore. Surface of DpsA (A) and DpsC (B) coloured according to the electrostatic surface potential. Red colour is negative; white is neutral and blue positive. Residues highlighted in DpsA at the "Dps like pore" are His-51, for DpsC Tyr-82, Gln-83 and Asp-194 are highlighted. Images were created using UCFS Chimera 1.7rc.

4.5.5.2 The internal cavity surface ESSP

The cavity of the protein can be explored in the same manner as the external surface. Visualisation of the electrostatic surface potential of the cavity reveals that it is not just the external surface of the dodecamer that employs electrostatic gradients to facilitate the movement of positive iron. Both DpsA and DpsC have cavities that have negative electro-potential surfaces. Additionally, differences in potential between the N-terminal and C-terminal pores are maintained, replicating what is found on the external surface. Moreover, the difference in potential between DpsA and DpsC is protracted within the cavity. Figure 4.14 demonstrates the negative electrostatic potential present on the inner surface of DpsA, around the N-terminal pore. When compared to DpsC, the negativity of DpsA's internal surface is massively extended which is visualised by a greater area coloured red. This pattern is recurrent when observing the potentials at the C-terminal pore and in addition the anti-parallel dimer complex. A noticeable trait present on the internal surface of the dimers is a "pathway" of negative electrostatic potential which runs from the region that creates the N-terminal pore, the proposed entrance for ferrous iron, diagonally across the four helix bundle and directly to the ferroxidase centres (FOC). In DpsC this trail of negative potential is created with Asp-167 and Asp-171, which are also the residues responsible for the negativity of the N-terminal pore. Likewise, residues from the "ferritin-like" pore of DpsA start the negative pathway once inside the cavity. Glu-140 and Asp141 start the route. However in addition to these, there are a further four acidic residues that extend the negative pathway from the N-terminal pore as far as the furthest FOC at the dimer interface of DpsA. Asp-72 is situated between the N-terminal pore and the closest FOC and helps to extend the negative gradient from Asp-141 thus traversing from the D helix over to the B helix. This "pathway" is continued via residues Asp-68 and Asp-64 located between the two FOC and finally Glu-57, which is beyond the furthest FOC. These negative gradients are indicative of a means of controlling the placement of iron once it is inside the protein. Unlike DpsC, iron that enters via one pore, can be transported to the furthest FOC possibly enhancing oxidation capabilities.

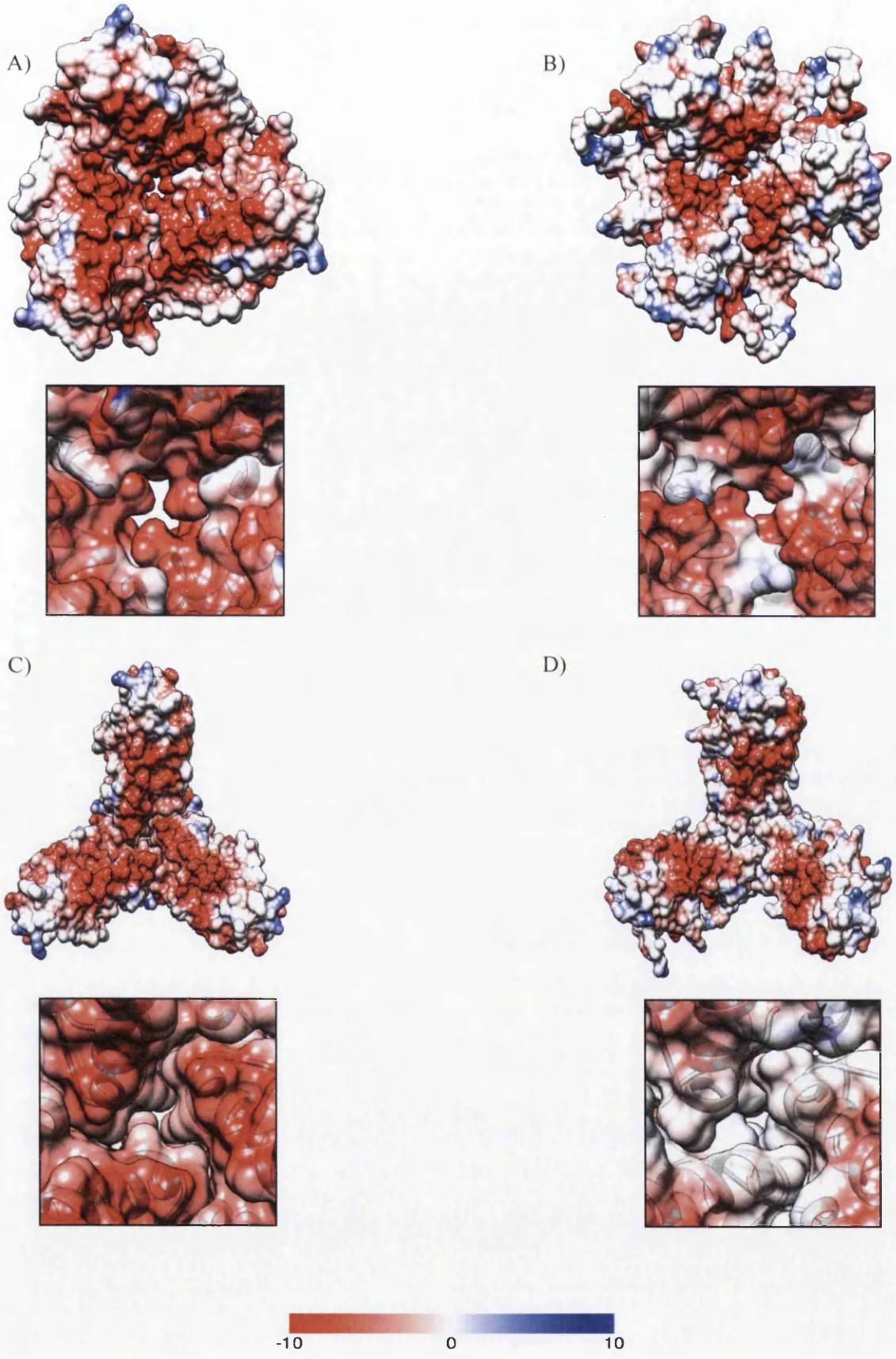


Figure 4.14: Electrostatic surface potential of the dodecamer cavities. Cavity surface of DpsA and DpsC coloured according to the electrostatic surface potential. Red colour is negative; white is neutral and blue positive. DpsA (A and C), Dps C (B and D). N-terminal pore (A and B) and C-terminal pore (C and D). Images were created using UCFS Chimera 1.7rc.

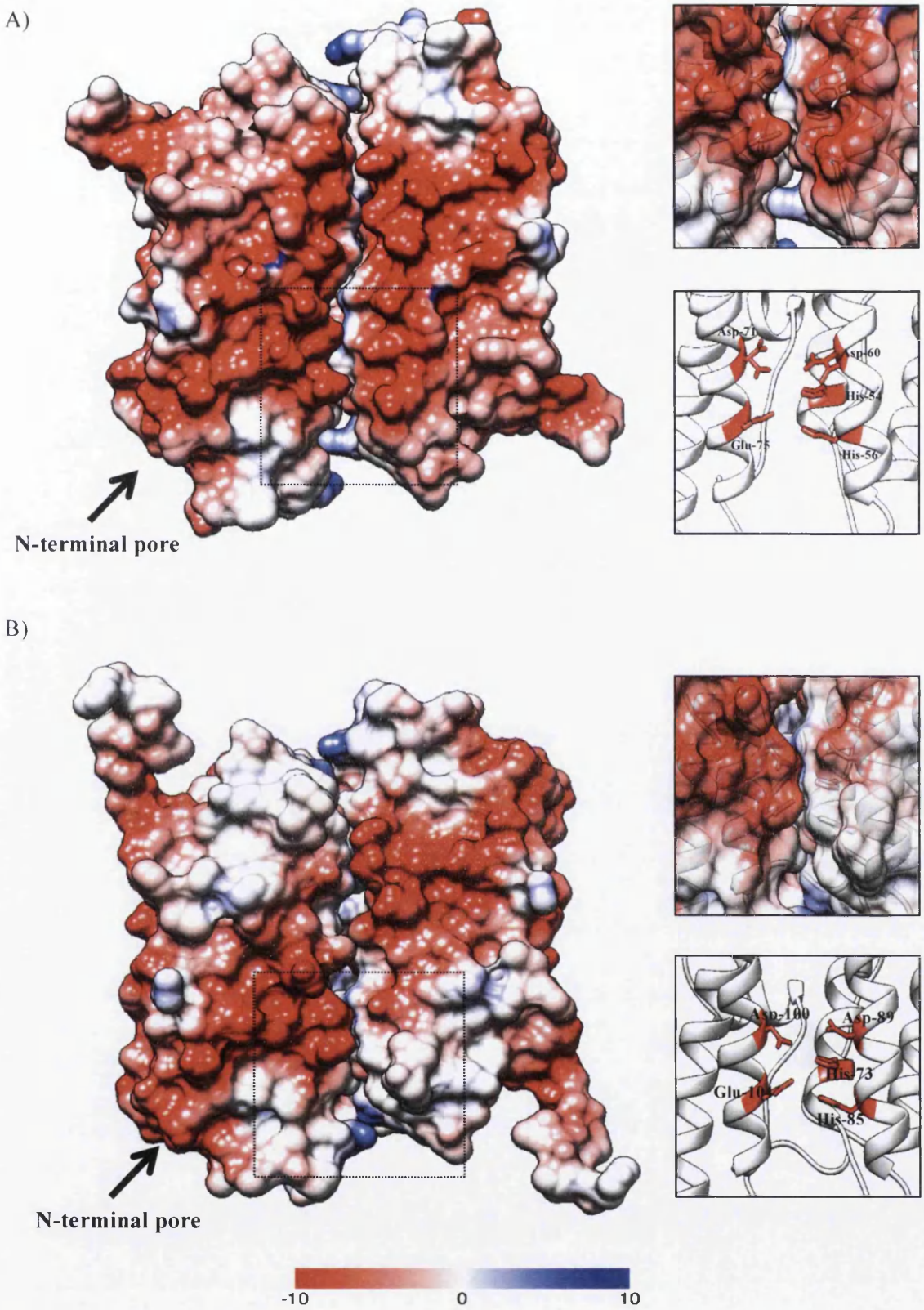


Figure 4.15: Electrostatic surface potential at the cavity side of dimers. Cavity surface of DpsA and DpsC coloured according to the electrostatic surface potential. Red colour is negative; white is neutral and blue positive. Antiparallel dimers with inserts highlighting residues donated to form ferroxidase centre. DpsA (A) donates residues His-44, His-56, Asp-60, Asp-71 and Glu-75. DpsC (B) donates residues His-73, His-85, Asp-89, Asp-100 and Glu-104. Images were created using UCFS Chimera 1.7rc.

4.5.6 Amino acid surface hydrophobicity molecular simulation

The hydrophobicity surface of the proteins, mapped to the Kyte-Doolittle hydrophobicity scale (Appendix 22) (Kyte and Doolittle, 1982), was created within UCSF Chimera and coloured blue for the most hydrophilic, to white, to orange for the most hydrophobic. The resulting images (Figures 4.16 and 4.17) highly resemble those of the electrostatic surface potential of the protein, with the N-terminal pore being hydrophilic in nature and the C-terminal pore hydrophobic in nature where in the case of the ESSP models instead there would be negative and positive surface potential. Furthermore the hydrophobicity of the internal surface is conducive to the transport of iron. A pathway of hydrophilic residues forms between the N-terminal pore and the FOC created at the dimer interface where a hydrophilic environment forms to allow for the coordination of iron via oxygen in order for oxidation to take place. A repetition of the ESSP models is once more revealed when comparing DpsA and DpsC. On the internal surface of the dimers, the hydrophobic pathway that dictates iron's route from the N-terminal pore to the ferroxidase centre is far greater in DpsA than DpsC where it would appear that iron may be able to access both FOC's in DpsA but is limited to the closest FOC to the N-terminal pore in DpsC. Together with the ESSP models of the cavity, this feature could pose a plausible explanation for the differences in ferroxidase activity observed between DpsA and DpsC in Chapter 3.

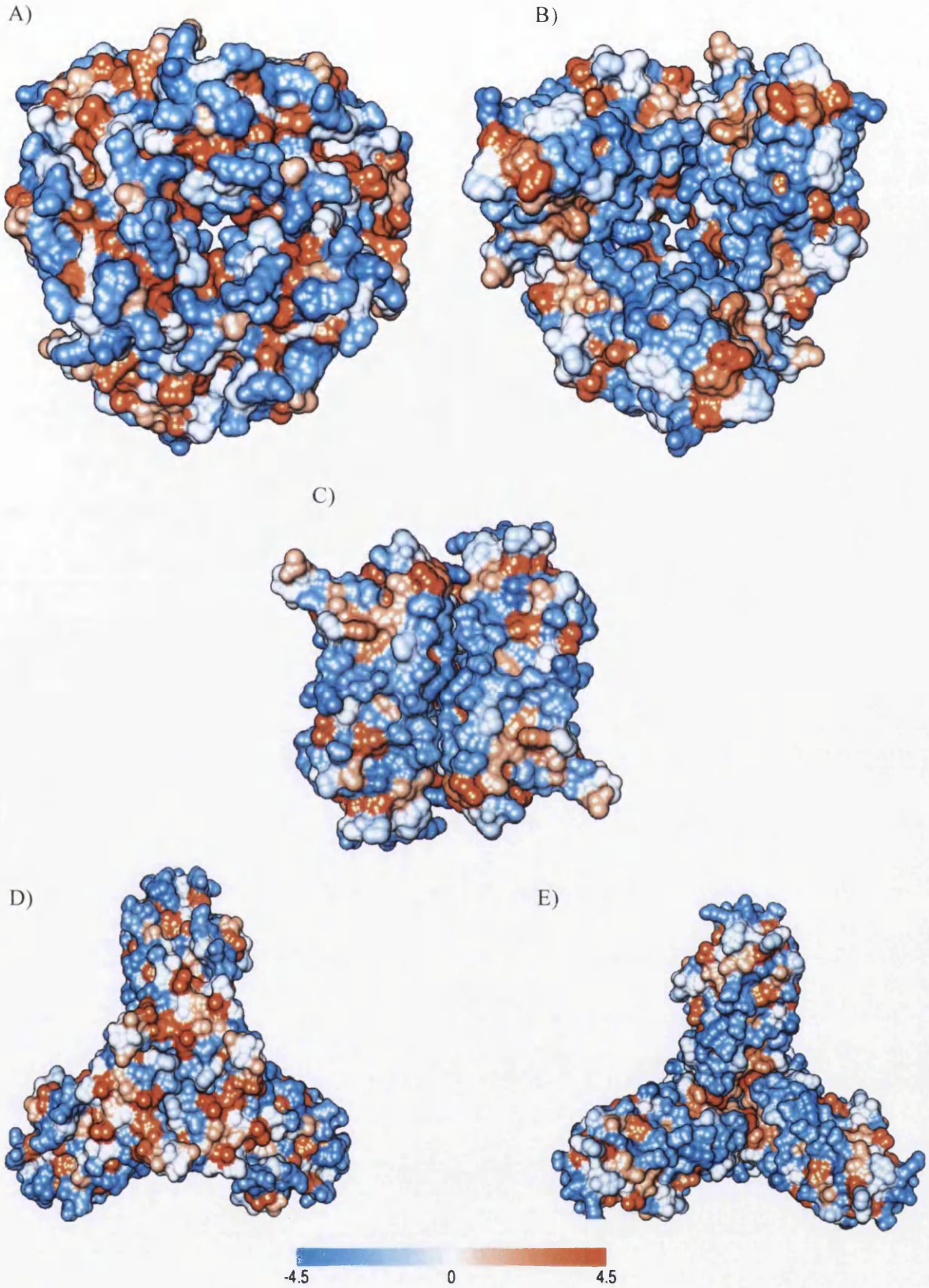


Figure 4.16: DpsA surface hydrophobicity coloured to the Kyte-Doolittle hydrophobicity scale. Blue is most hydrophilic and orange hydrophobic with white in between. External surface at N-terminal trimer pore (A) and C-terminal trimer pore (D). Internal surface at N-terminal pore (B) and C-terminal pore (E) with the internal surface of the antiparallel dimer (C). Images were created using UCFS Chimera 1.7rc

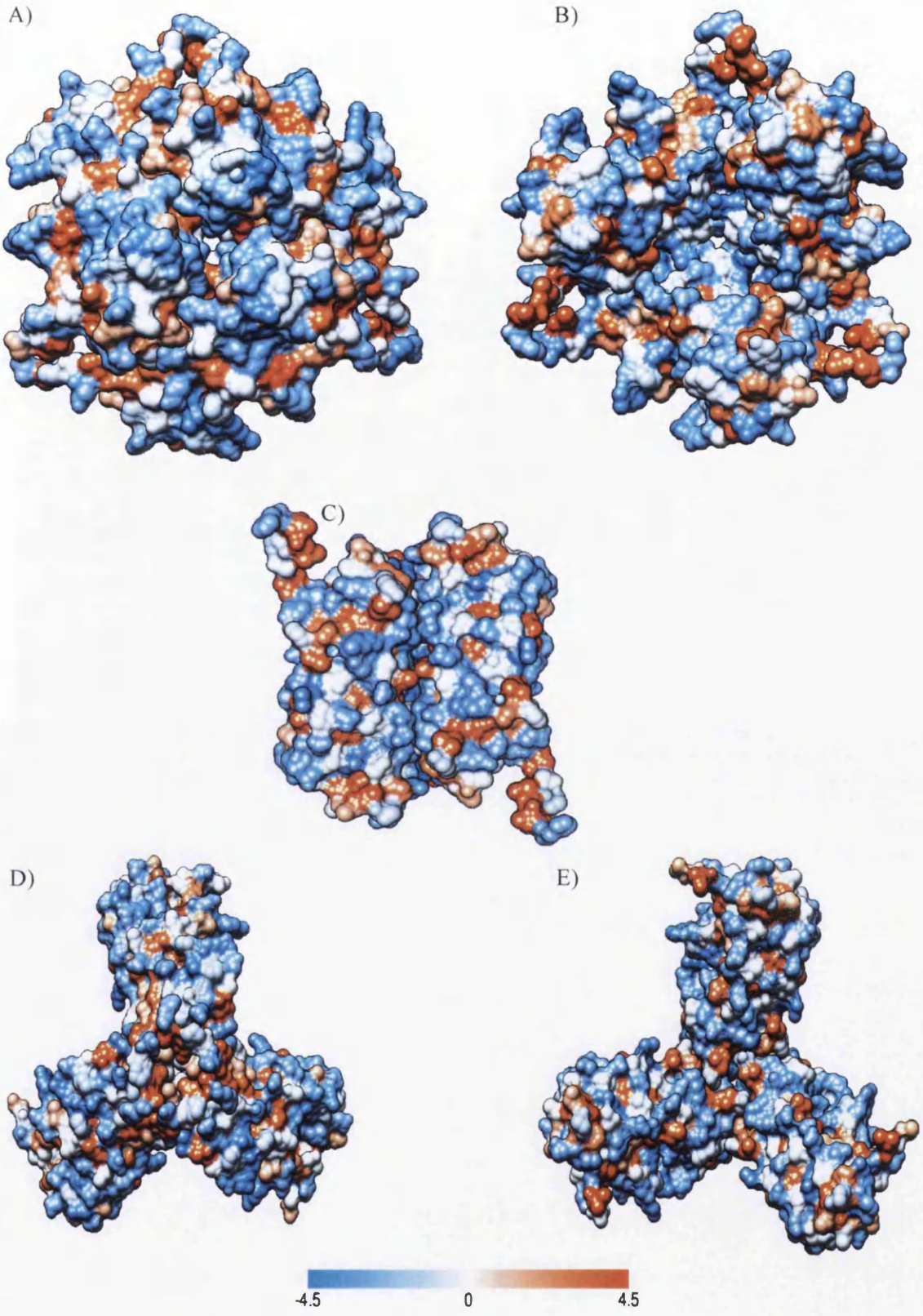


Figure 4.17: DpsC surface hydrophobicity coloured to the Kyte-Doolittle hydrophobicity scale. Blue is most hydrophilic and orange hydrophobic with white in between. External surface at N-terminal trimer pore (A) and C-terminal trimer pore (D). Internal surface at N-terminal pore (B) and C-terminal pore (E) with the internal surface of the antiparallel dimer (C). Images were created using UCFS Chimera 1.7rc

4.6 Summary of Resolution and Analysis of the Crystal Structures of *Sc*DpsA and *Sc*DpsC results

Resolving and analysing the crystal structures of DpsA and DpsC has revealed the structural features that give rise to the biochemical differences found between the two proteins in the previous chapter. Not only does this provide insight into how the proteins are adapted for their *in vivo* roles, but the anaylsis also provides a basis from which to continue exploring the proteins in order to develop them as biotechnological tools. PISA and virtual alanine scanning have provided a list of residues that are energetically significant to how the proteins self-assembly and in the case of DpsC, remain highly stable. Thus the summary of results that follows not only gives an overview of this chapter's main findings, but also highlights the potential future work that can be carried out.

- Both DpsA and DpsC display stereotypical four helix bundle monomers that form dodecameric assemblies with 233 point symmetry.
- The tail regions display some disorder within the structures; 18 residues of the native DpsA C-terminal tail and 27 residues of the DpsC N-terminal tail are disordered.
- The superimposed monomeric structures of DpsA and DpsC yield a RMSD value of 1.047 Å, with DpsA structurally similar to the *Mycobacterium smegmatis* Dps (RMSD: 0.681 Å) and DpsC structurally similar to the *Thermosynechococcus elongatus* Dps (RMSD: 0.761 Å). Largest deviations occur within loop and tail structures.
- The external diameters of the dodecamers are in agreement with those reported for other Dps measuring ~87 Å in diameter with the protein cage creating internal cavities for the deposition of ferric iron.
- Analysis of protein-protein interfaces within the structures (PISA) display significant differences between DpsA and DpsC.
    - DpsC has an extended dimer interface area (1870.5 Å), the highest of all the dimer interfaces explored, even Dps from extreme environments.
    - DpsC dimer interface is stabilised by numerous hydrophilic interactions (27 hydrogen bonds and 27 salt bridges), a total of 32 more than those across the interface of DpsA.
    - The N-terminal trimer "ferritin-like" interfaces are similar in the two proteins, but the C-terminal trimer "Dps-like" interface has a greater area, is vastly more hydrophobic and more significant to assembly in DpsC.

- Analyses indicate both N and C-terminal tails are involved in multiple interfaces often involving hydrophilic interactions, implicating the tails in a role of maintaining and stabilising the dodecameric assemblies.
- The interface statistics of DpsC make it equivalent to those from extremophilic-like Dps, while DpsA ranks among mesophilic-like Dps.

- Virtual alanine scanning predicts numerous hotspot residues to be energetically significant to the interface in both DpsA and DpsC.
  - Hotspot residues are predicted at each of the interfaces of DpsA and DpsC.
  - DpsA has predicted hotspots within the N-tail that take part in the "ferritin-like" interface.
  - DpsC has predicted hot spots in the N-tail at the dimer interface, and "ferritin-like" interface, in addition the C-tail has hotspots important to the "Dps-like" interface.
  - The extremely hydrophobic "Dps-like" interface of DpsC has 4 aromatic residue hotspots out of a total of 9 hotspots.
  - The most significant hotspot of DpsC is located on the BC helix (Arg-118) alluding to a role for this helix in conferring stability.
- Electrostatic surface potential and hydrophobicity profiles demonstrate the hydrophilic, negative nature of the "ferritin-like" pores of DpsA and DpsC.
- The "Dps-like" pore is hydrophobic and has vastly reduced negative electrostatic potential. His-51 residues of DpsA not only block the channel but can neutralise any negative gradients. This implies the channel is not likely to facilitate the transport of large cations such as iron.

Chapter 5

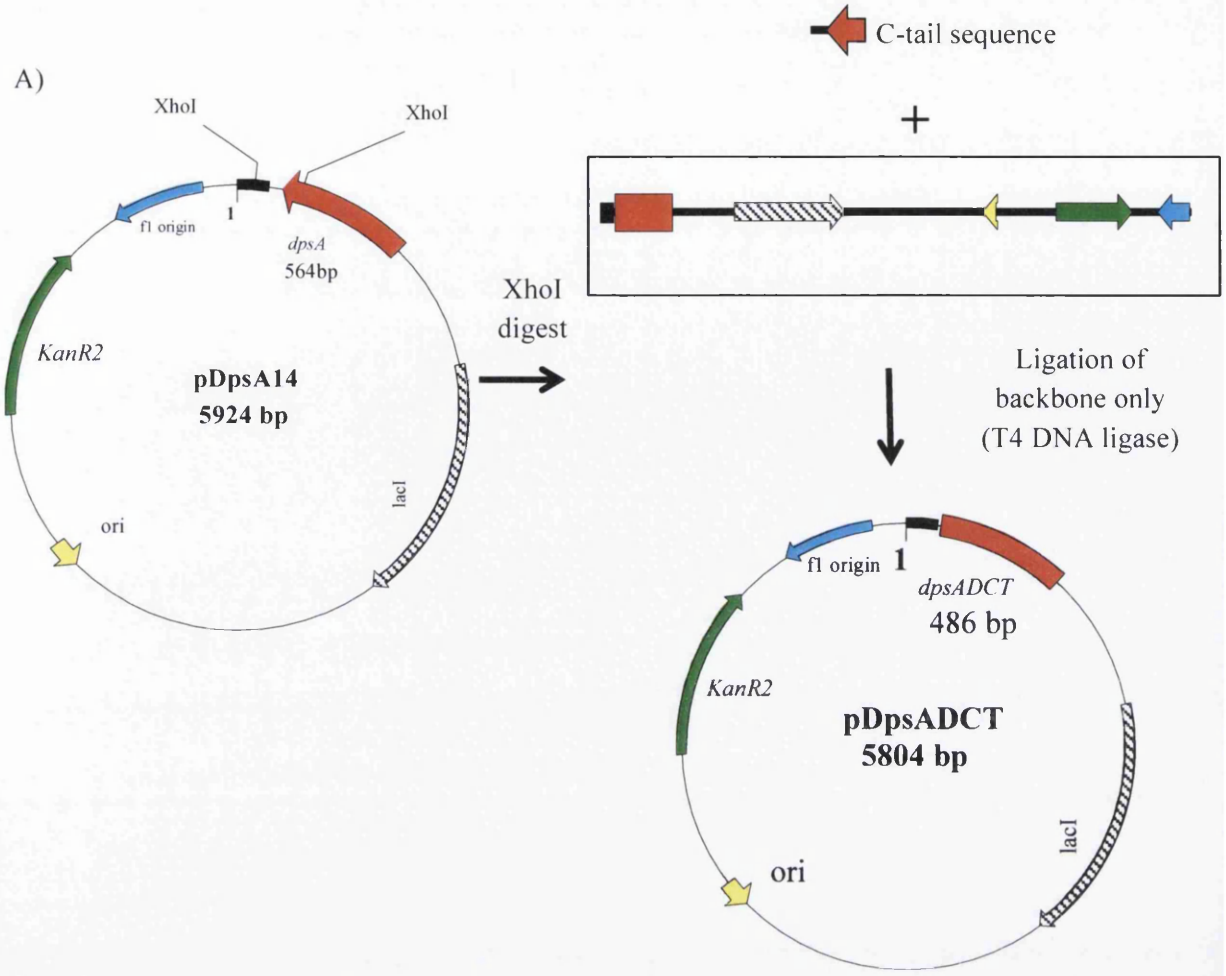
Protein Structure Manipulation and Analysis of Mutant Protein Structure

5.1 Protein Manipulation

This chapter describes the manipulation of protein sequence in order to assess the influence that the altered regions have on the structure/function relationship so often discussed when studying proteins. It has been reported that the N and C tail extensions that protrude from the core four helix bundle of a Dps monomer are involved in oligomeric assembly and stability. The previous chapter has already highlighted how two of the Dps proteins from *S. coelicolor* (DpsA and DpsC) are capable of assembling into typical Dps dodecamers, however, by simply tagging DpsA at the N-terminal it was possible to distort the assembly and stability of the complex, resulting in the formation of soluble aggregates. This highlighted how the N-tail of DpsA may be involved in oligomeric assembly. Furthermore, the crystal structures revealed interactions between the terminal tails themselves and between each tail and the four-helix bundle. This confirms the notion that the tails are involved in assembly and stability. However, there is only one way of truly assessing the contribution of the N and C tails to assembly and that is to remove these regions and monitor how this affects the protein's biochemical properties.

5.1.1 Creation of Tail-less Dps Mutant Proteins

Three DpsA mutant constructs were created which coded for proteins lacking the N-terminal tail of 14 residues (pDpsADNT), one lacking a 25 residue C-terminal tail (pDpsADCT) and a protein lacking both tails and a total of 39 residues deleted (pDpsADTM). A single mutant construct was created for DpsC, which lacked the 44 residues, which compose the N-terminal tail (pDpsCDNT). All plasmid maps are displayed in Figure 5.2. The genetic constructs of these mutants were created in two ways. All but pDpsADCT were created by designing mutant sequences coding for the proteins lacking the tails, these sequences were then codon optimised for expression in *E. coli*, the genes were commercially synthesised by Eurofins and cloned into pCR®2.1 plasmids for further sub-cloning. Each gene was cut out of the pCR®2.1 plasmids using *NdeI/Bgl*II and ligated to pET-26b (+) plasmids, pre-cut with *Nde*I/*BamH*I. Successful cloning was confirmed by sequencing using the T7 forward primer and T7 reverse primer. DpsADCT was created utilising the existing *Xho*I cut site within the *dpsA* sequence and the *Xho*I cut site within the multi-cloning site of the pET-26b (+), more specifically within the pDpsA14 vector. Plasmid pDpsA14 was cut using *Xho*I to release a small 120 bp fragment coding for the C-tail and linker region. The plasmid backbone containing the newly created C-tail-less DpsA gene was extracted from an agarose gel and re-ligated to form pDpsADCT. The *Xho*I re-ligation preserved the correct reading frame, ensuring that the C-terminal Histidine tag was maintained (Figure 5.1A). This construct was also confirmed by sequencing.



B)

A

MTHDLTPKYTVPGIEREAAGRLIGVLRLRLHALNDLHLTLKHVHWNVVGPH

B **BC**

FIAVHEMIDPQVDQVRDMADDVAERIAALGGVAQGTPGALVAERKWDDYS

C **D**

IGRADAIAHLGALDVVYTGVVEGMRAAVEEAGKIDPATEDLLIGQLRDLEQF

QWFVRAHLESAGGALATGSATSETEAAARGAEVADPNSSSVDKLAAALEHHH

HHH

Figure 5.1: Cloning steps involved in the construction of DpsADCT. (A) The *Xho*I restriction site within the sequence of *dpsA* was used along with the *Xho*I site present in the MCS of the pET-26b (+) backbone to remove the sequence coding for the C-terminal tail of DpsA along with the linker region of between protein and His tag. The linear plasmid was purified and re-ligated to form pDpsADCT. Maps not drawn to scale. B) Protein sequence of DpsA with residues underlined signifying those deleted for deleted tail mutant constructs; helices are shaded in grey and the helix labelled above.

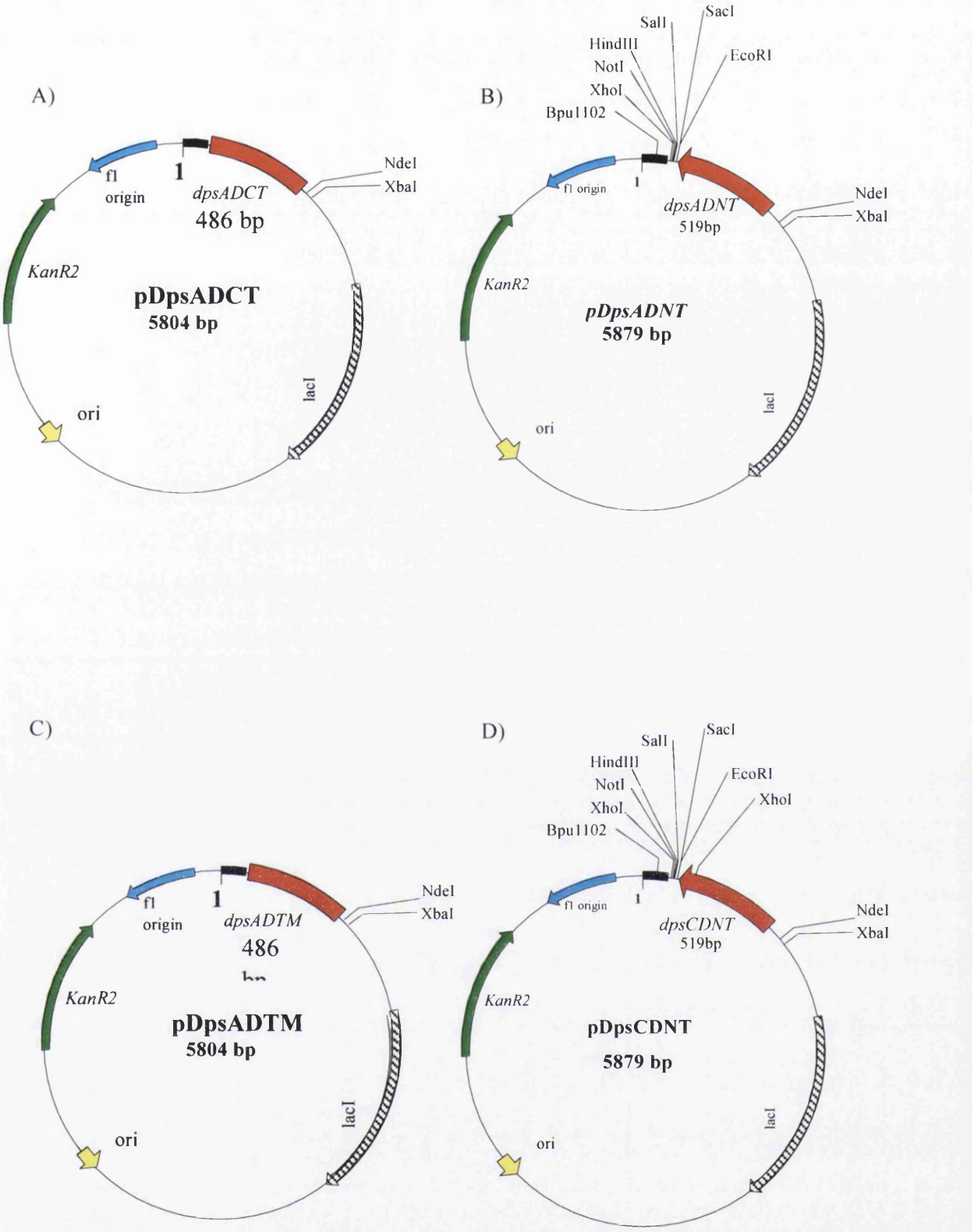


Figure 5.2: Protein expression plasmid maps of truncated DpsA and DpsC constructs. DpsADCT (A) was cloned by removing a fragment coding for the C-terminal tail by restriction digest via *Xho*I and relegation. DpsADNT (B), DpsADTM (C) and DpsCDNT (D) were all synthetic gene constructs and were cloned into pET-26b (+)

5.1.2 Recombinant Expression of Dps Tail Mutants

The recombinant expression of the mutant proteins was achieved using the same protocol for the full length proteins, where growth of the cultures was carried out at 37 °C until and OD at 600 nm reached 0.9 before induction of protein synthesis was carried out at 30 °C. The mutant proteins were purified from the soluble fraction of the BL 21(DE3) cells as described in previous chapters although noticeably lower protein concentrations were observed.

5.2 Size Analysis of Mutant Proteins

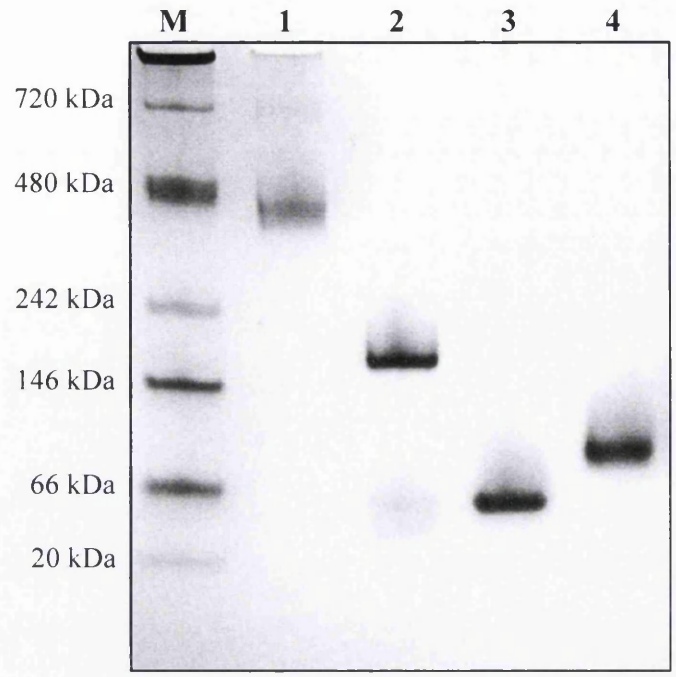
5.2.1 DpsA Deleted Tail Mutants – Removal of the Tails Disrupts Oligomeric Assembly

Once purified, the proteins were subjected to native PAGE in order to predict their molecular weight and thus calculate their oligomeric assembly. Figure 5.3A shows the native PAGE gel stained with Coomassie R-250, the results indicating that removal of any tail severely disrupts oligomeric assembly of DpsA. Having both lengthy N and C-terminal tails, DpsA was manipulated to create three mutants; deleted N-tail (DpsADNT), deleted C-tail (DpsADCT) and both N and C-tails (DpsADTM). All these truncated proteins failed to assemble as the full length protein does in solution. DpsADCT migrates with a molecular mass just above that of the 146 kDa maker and is the highest oligomeric state of the mutant species. DpsADNT migrates above the BSA marker of 66 kDa whilst DpsADTM migrates roughly in line with this BSA marker. Using the molecular weight ladder it is possible to create a standard curve of migration against log of the molecular weight. This can be used to predict the molecular weight of the samples based on their very own migration using the equation of the linear trendline. Table 5.1 documents the molecular weights predicted for each protein run on the native PAGE gel displayed as Figure 5.3A and predicts their oligomeric state.

Table 5.1 Molecular weight predictions via migration of standards on Native PAGE

Protein	Migration (mm)	Log Mw	Mw	Predicted monomeric Mw	Oligomeric state
DpsA	35	2.587752	387.03	22.37	16-mer
DpsADCT	64	2.192904	155.92	18.68	8-mer
DpsADNT	83	1.93421	85.94	20.82	4-mer
DpsADTM	93	1.798056	62.81	18.63	3-mer

A)



B)

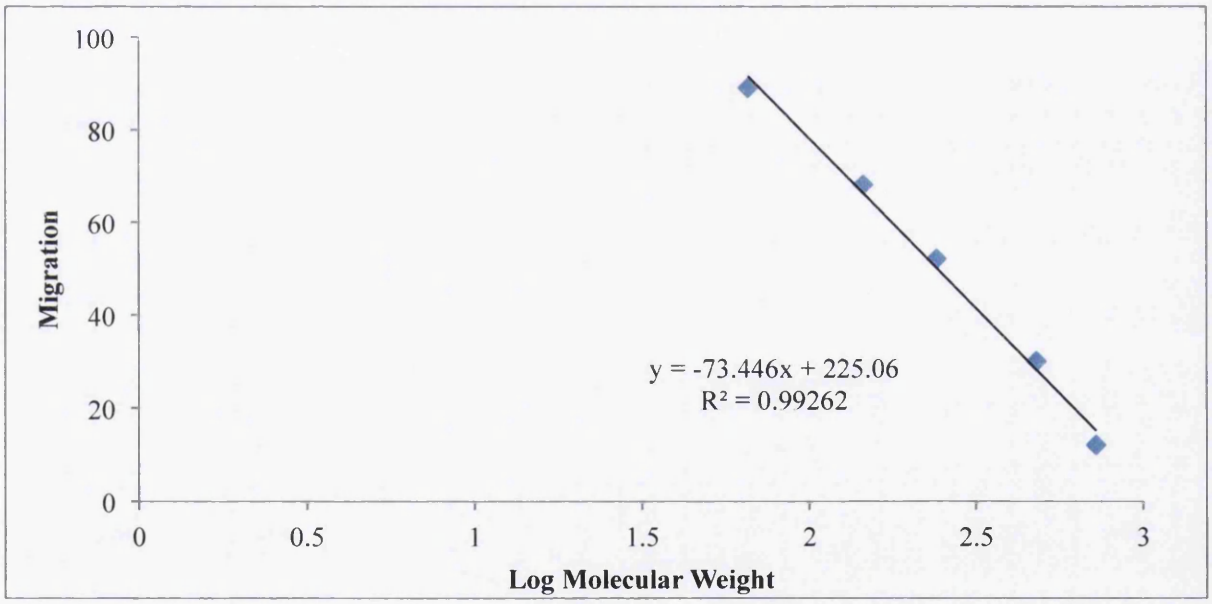


Figure 5.3: Size analysis of DpsA mutant proteins lacking terminal tails. (A) Native PAGE gel stained with coomassie R-250 of DpsA (lane 1), DpsADCT (lane 2), DpsADTM (lane 3) and DpsADNT (lane 4). The marker (lane M) was also used to calculate the molecular weight of the mutant species via a standard curve of migration distance plotted against log molecular weight. The points on graph (B) correspond to Apoferritin minor band (770 kDa), Apoferritin major band (480 kDa), B-phycoerythrin (242 kDa), Lactate Dehydrogenase (146 kDa) and Bovine Serum Albumin (66 kDa). Soybean Trypsin Inhibitor (20 kDa) was not included in the size analysis as it was outside of the linear range.

Despite the size estimations being very much an approximation given prediction of DpsA as a 16-mer, which is unlikely, the impact the tails have on assembly is immediately noticeable from the gel and the accompanying statistics. In accordance to the data presented in Chapter 3, the effect the N-terminal tag has on DpsA assembly is somewhat corroborated by the disruption to oligomeric assembly caused by the deletion of the N-terminal tail, which reduces oligomeric assembly to an estimated tetrameric state. The C-terminal tail is also shown to play a role in oligomeric assembly, although to a lesser degree. A predicted octamer is formed by DpsADCT, an assembly which is not too dissimilar to the *M. smegmatis* deleted C-tail mutant Dps (Roy et al., 2007), in which 26 residues are removed from the C-terminal causing the formation of a decamer (10-mer).

5.2.2 The DpsCDNT mutant forms soluble aggregates with high molecular weight

DpsCDNT, with a substantial 44 amino acid residues removed from the N-terminal (Figure 5.4A), was purified just like the DpsA tail mutant proteins from the soluble extract of induced *E. coli* BL21 (DE3) cells. However when resolving the fully soluble protein through blue native PAGE, the Coomassie R-250 staining revealed no dodecameric assembly present in the gel (Figure 5.4B). Instead, protein was detected within the wells of the gel with a small smear effect below each well suggesting that whilst the protein was fully soluble, it was suffering from non-specific aggregation. Although steps were taken to remove insoluble matter via syringe filtration (0.2 μM pore size), the inability to migrate through the gel persisted. Despite the fact that native DpsC cannot be denatured with 8 M urea, DpsCDNT was placed in 8 M urea for 24 hours, the urea was then removed and the protein desalted into a range of NaCl concentrations from 200 mM to 500 mM, all buffered with Tris/HCl pH 7.5 containing 5% glycerol (v/v). The varying salt concentrations were used to cover a range of ionic strengths in order to find a condition suitable for any protein refolding that may occur. As for DpsC, its deleted N-tail mutant also resists denaturation via urea. As a result of the inability to be denatured, the refolding steps had little impact on the assembly or aggregation of DpsCDNT although less protein can be seen entering the gel at higher ionic strengths.

A)

MSSPKPKPPSSAEHRSDGSQPWLHQKGRTIQEFGTVKQFPVALTMDTRLYSC

A B

QRLNKVLADTRILHDLYKKYHWLMRGATFYQLHLLLDKHAGEQLELIDTV

BC C

AERVQTLGGVAVGDPRHVAEITTVPRPPDGVEEVPSMLSRLLEAHELILTEC

D

HDAAARTQEYGDDGTNDLLVSEVLRTNELQAWFVAEHLVDTPLVHADPNS

SSVDKLAAALEHHHHHH

B)

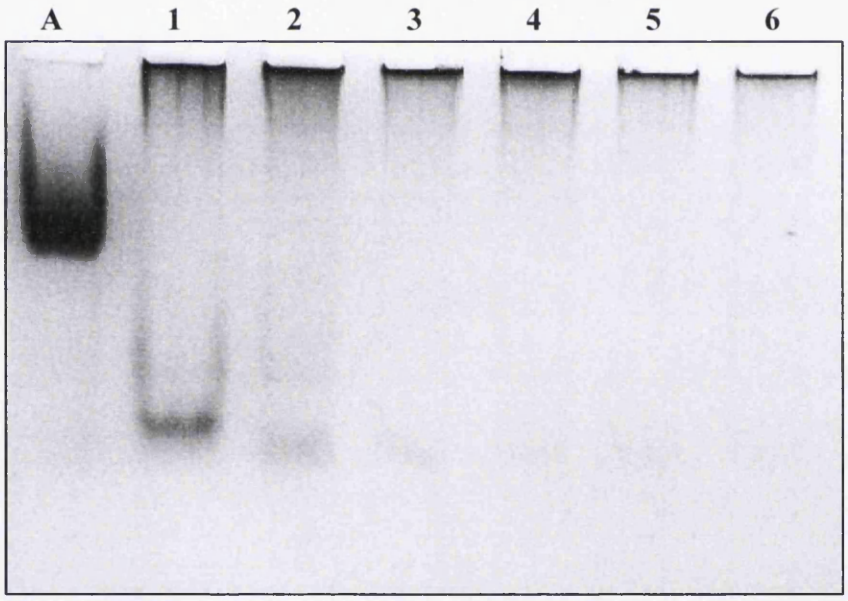


Figure 5.4: Analysis of DpsADNT sequence and oligomeric assembly. (A) Protein sequence of DpsC with the 44 residues deleted to create DpsCDNT underlined. The helices are shaded in grey and are labelled above the sequence. (B) Native PAGE gel of DpsCDNT before addition of urea, with urea and after the removal of urea into different salt concentrations (all solutions were buffered with 20 mM Tris-HCl pH 7.4 and contained 5 % glycerol). Lane A – DpsA, lane 1 – DpsCDNT, lane 2 – DpsCDNT 8 M Urea, lane 3 – DpsCDNT in 200 mM NaCl, lane 4 – DpsCDNT in 300 mM NaCl, lane 5 – DpsCDNT in 400 mM NaCl, lane 6 – DpsCDNT in 500 mM NaCl. The lower band of lane 1 may be indicative of partial dissociation or degradation.

5.3 Ferroxidase Activity of Tail-less Mutant Constructs

5.3.1 Ferroxidase Activity Is Maintained in the DpsA Tail-less Mutants

For ferroxidase activity to occur, Dps proteins require the formation of the ferroxidase centre at the interface between two antiparallel subunits, so preservation of ferroxidase activity would therefore confirm the conservation of a dimer assembly. Since all the deleted tail mutants of DpsA form multimeric assemblies, each mutant was tested for the ability to oxidise ferrous iron. As described in Chapter 3 (3.3), ferroxidase activity was monitored in the presence of oxygen as an oxidant. A single condition was used in order to compare these mutant variants with the full-length protein. Figure 3.4 demonstrates that 276 μM ammonium iron (II) sulphate is a concentration where the initial oxidation rate cannot be increased by adding more iron, thus this concentration was used to study iron oxidation in the DpsA mutant variants.

All the mutant variants had ferroxidase activity, and even DpsADTM, predicted to be a trimer, retained ferroxidase activity. However surprising oxidation curves were produced showing that all the mutant proteins have disparate reaction rates, and to some degree differed in their ability to oxidise increased amounts of iron during the reaction time. DpsADCT shows a marked increase in initial ferric iron formation rate and has an enhanced final absorbance when compared to DpsA at 1200 seconds (0.623 compared to 0.512). The DpsADCT progression of oxidation levels off at 190 seconds with an absorbance of 0.626 at 310 nm. At 190 seconds the absorbance caused by ferric iron within the DpsA reaction is 0.317, not far from half of the absorbance caused by DpsADCT. By the end of the reaction the absorbance of DpsADCT drops to 0.623 at 1200 seconds.

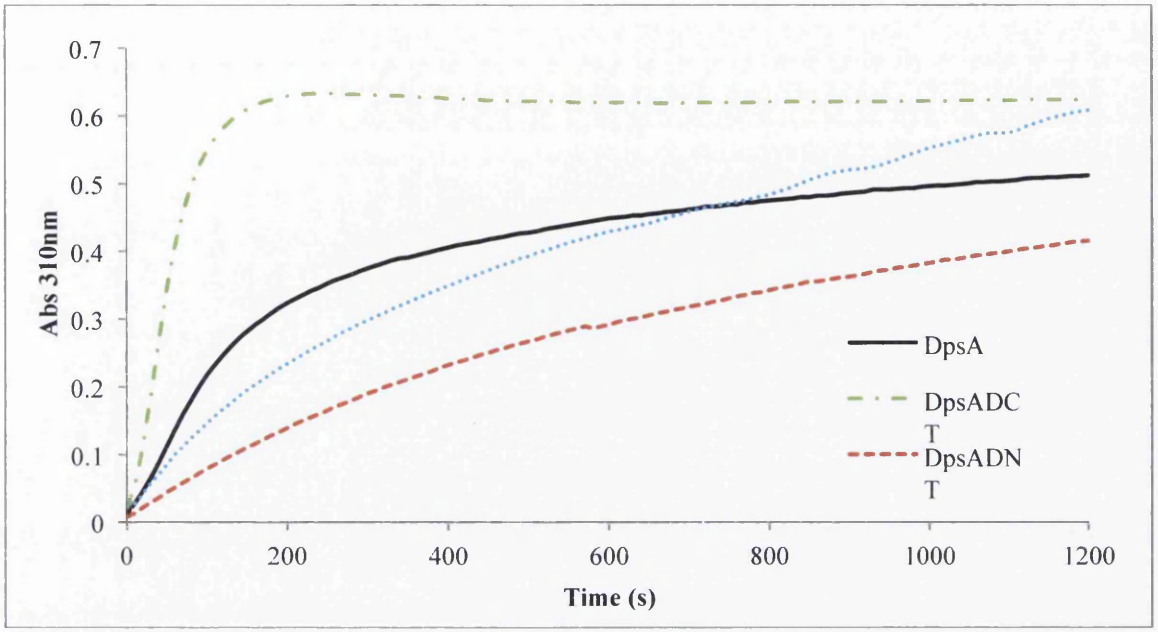
DpsADNT had the slowest progression curve of the mutant proteins and yielded the lowest absorbance (0.415) by the end of the reaction. Perhaps the most unexpected curve was that produced by DpsADTM. The smallest of all the oligomeric assemblies, DpsADTM produces a final absorbance above DpsA and more equal to DpsADCT, a result unexpected given that a trimer assembly is unlikely to create a dimer interface and thus a FOC. Whilst having a curve lower than DpsA, indicating a slower initial rate, the absorbance continues to climb and does not appear to level off as it does for DpsA and DpsADCT. The ability of these proteins to mineralise and bind the oxidised iron within these reactions was tested and the results follow in the next section.

5.3.2 Ferric Iron Mineralisation

To determine if the mutant proteins were capable of binding the ferric iron, *in vitro* in gel ferric iron staining was employed. Chapter 3.3.3 describes how both DpsA and DpsC were capable of depositing iron within their internal cavity. Ferroxidase activity reactions used for spectrophotometric analysis were incubated at room temperature for 1 hour to allow completion of oxidation and full mineralisation to occur. Protein solutions were then processed for native PAGE; insoluble material was pelleted by centrifugation and the soluble content of the reactions maintained in the supernatant. The proteins were placed in fresh iron free buffer with the use of size exclusion chromatography (HiTrap Desalting columns) and were concentrated to 2.0 mg/ml using centrifugal filtration devices with a 20 kDa molecular weight cut off. At this stage it was possible to visualise a slight "rusty" colour present within the DpsADCT solution, an indication of the presence of iron. Neither DpsADNT nor DpsADTM retained this colour but instead were colourless clear solutions. Following the concentration stage, the proteins were resolved by clear native PAGE before the iron staining procedure was carried out. Iron loaded DpsA was also used as a positive control.

Interestingly, of the three mutant species, only one was capable of binding ferric iron (Figure 5.5B). DpsADCT gave a positive royal blue band in the iron staining procedure whilst DpsADNT and DpsADTM failed to give a signal. Coomassie R-250 staining established that both DpsADNT and DpsADTM were present in the gel and that the presence of iron does not alter their oligomeric assembly confirmed the presence of protein. The overlapping band pattern of Coomassie and ferric iron staining demonstrates that the iron is only present in the higher oligomeric assembly of DpsADCT. This suggests that there is some form of cavity present within the DpsADCT assembly, which is capable of holding the mineralised ferric iron nano-core. Conversely DpsADNT and DpsADTM are small oligomeric species, potentially of tetrameric and trimeric nature respectively, and thus would not allow the formation of a cavity space.

A)



B)

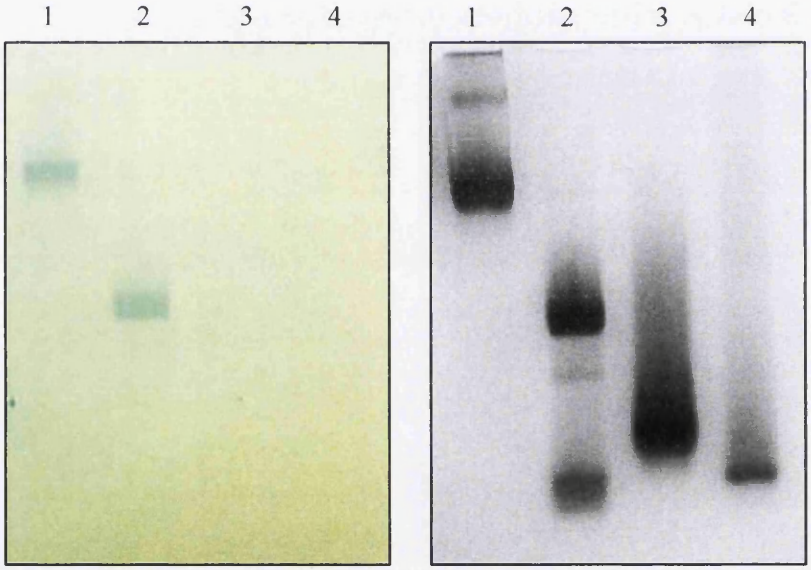


Figure 5.5: Ferroxidase activity and iron binding of DpsA tail-less mutant variants. (A) Ferroxidase activity of DpsA (black) and its tail mutants, DpsADCT (green), DpsADNT (red) and DpsADTM (blue). Each curve represents three technical replicates. (B) native-PAGE in gel potassium ferricyanide ferric iron staining gives a positive royal blue band result for iron loaded DpsA (lane 1) and DpsADCT (lane 2). DpsADNT (lane 3) and DpsADTM (lane 4) fail to show iron binding ability even though they retain ferroxidase activity.

5.3.3 DNA Oxidative Protection and DNA Interaction

In chapter 3 it was described how hydrogen peroxide and ferrous iron were the causative agents of *in vitro* DNA damage as a result of the generation of harmful ROS via Fenton chemistry. However, in the presence of DpsA, complete degradation of DNA was circumvented by the iron and hydrogen peroxide scavenging nature of the Dps protein. Since the DpsA mutant proteins retained ferroxidase activity, their ability to protect DNA against oxidative damage was also investigated. Analysis of the ferroxidase activity of tail-less DpsA was carried out using oxygen as an oxidant; hydrogen peroxide was not tested. Thus this experiment was used as a means of not only establishing if loss of tails influenced the ability of the proteins to protect DNA *in vitro*, but also provides an opportunity to assess the capability of the protein to utilise hydrogen peroxide. The reaction conditions employed were as used for DpsA in chapter 3.

Figure 5.6 depicts the results of the *in vitro* DNA protection assay in which DNA was exposed to Fenton reagents in absence and presence of the proteins. When DNA is treated with either ferrous iron or hydrogen peroxide little effect is observed on the integrity of the DNA as visualised by the preservation of the DNA band. The combined effect of both iron and H_2O_2 with DNA results in the band no longer being visible within the lane; this is indicative of complete DNA degradation. All tail mutant variants conferred protection of the DNA by means of their iron oxidation ability. Whilst quantitative analysis was not performed, it can be suggested that the DpsA tailed mutants afford marginally weakened oxidative protection, which is exhibited by slightly diminished intensities from the DNA bands when compared to DpsA.

DpsA did not demonstrate any form of DNA interaction in any of the conditions tested with or without the addition of a divalent metal ion in chapter 3. Accordingly it is unlikely that any of the mutant variants would reverse this behaviour, and indeed no DNA interaction was detected when analysing the DNA oxidative protection assay.

Lane	1	2	3	4	5	6	7	8	9	10	11	12
Protein	-	-	-	-	A	ADCT	ADNT	ADTM	A	ADCT	ADNT	ADTM
H_2O_2	-	+	-	+	-	-	-	-	+	+	+	+
Fe (II)	-	-	+	+	-	-	-	-	+	+	+	+

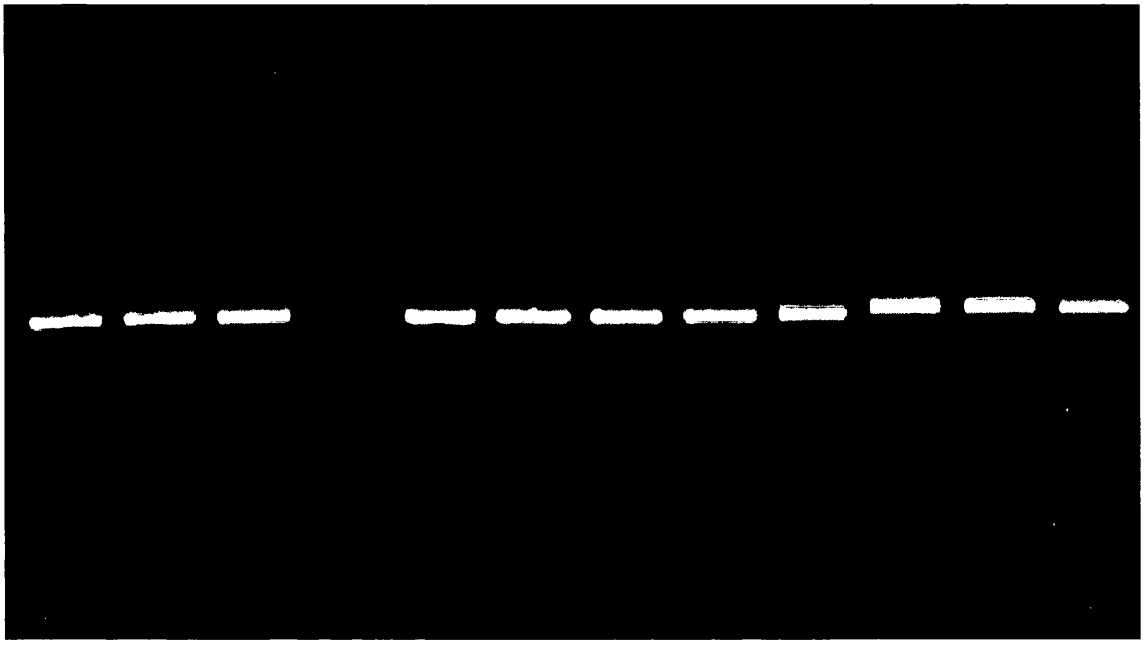


Figure 5.6: DNA oxidative protection by DpsA tail-less mutant variants. DpsA and its deleted-tail variants were capable of protecting DNA against damage by ROS produced by Fenton chemistry in a reaction involving ferrous iron and hydrogen peroxide. All reaction mixtures contained 0.022 pmol linear pUC18. The table above the gel describes the conditions in which the DNA was incubated prior to resolving via agarose electrophoresis (1 %) and stained with ethidium bromide.

5.4 Generation of DpsA deleted BC helix

Another small but interesting feature of the Dps proteins that separates them from conventional ferritins is the placement of the small fifth helix. In the case of classical ferritins and bacterial ferritins, the E helix is positioned at the end of the four helix bundle (Khare et al., 2011), however the small helix in Dps, the BC helix, is positioned at the centre and perpendicular to the length of the four helix bundle. The BC helical residues also take part in the formation of the dimer interface. Prior to the determination of the DpsA crystal structure, a mutant construct was created in which the BC helix was deleted completely. A synthetic gene encoding DpsAdelBC was cloned from a pCR®2.1 plasmid (Eurofins) by using an NdeI/BglII double digest and subsequently ligated into a pET-26b$^+$ plasmid pre-cut with NdeI/BamHI. Cloning was confirmed via restriction and sequencing. Expression of the protein was achieved

as described for the other mutant proteins, with very little problem with insoluble precipitation and the protein was purified to homogeneity via IMAC using the hexa-histidine C-terminal tag.

Oligomeric assembly of the DpsAdelBC species was analysed via blue Native PAGE. DpsA was used as a control. Figure 5.7 displays the native PAGE gel stained with coomassie R-250 which indicates loss of the BC helix prevents dodecameric assembly. Whilst a predominant assembly with a molecular weight above 66 kDa but below 146 kDa was visible, there was also the presence of other lower abundance bands of greater molecular weight. A ladder effect was created which indicates multiple oligomeric assemblies, yet none of these bands migrated as full length DpsA did. In order to ascertain if tertiary level folding had any effect on the ability of the mutant protein to assemble, 8 M urea was used to disrupt quaternary structure and was then subsequently removed to allow protein refolding to occur. This had little impact on DpsAdelBC, which still failed to assemble and migrate as the full-length protein does. This finding highlights the multifaceted approach the Dps proteins use to maintain assembly, as no single structural element is uniquely responsible for mediating assembly.

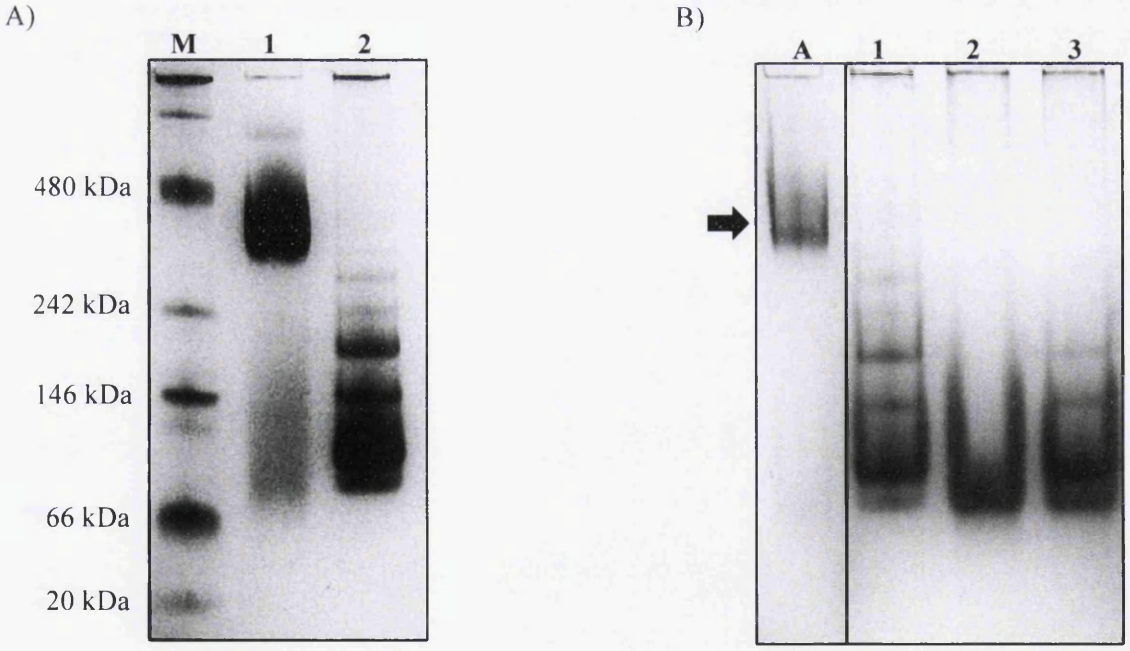


Figure 5.7: Size analysis of DpsAdelBC helix. (A) A non-dodecameric multi species assembly is observed in a Coomassie stained BN-PAGE; lane M – marker, lane 1 – DpsA and lane 2 – DpsAdelBC. Refolding of DpsAdelBC failed to improve oligomeric assembly (B). Composite image taken from same PAGE gel; Lane A – DpsA, Lane 1 – DpsAdelBC, lane 2 – DpsAdelBC 8 M Urea and lane 3 – DpsAdelBC refolded from 8 M Urea. The arrow indicates the would-be-migration of a dodecameric assembly

5.5 Crystal structure of DpsADCT

DpsA has a C-terminal tail that is 26 amino acid residues in length. In the crystal structure of DpsA 18 of these amino acids were in a disordered state and accordingly were not visible. However experimental evidence has shown how the deletion of the C-terminal tail of DpsA results in the abolition of dodecameric assembly. Furthermore a different and unusual multimeric assembly was predicted to occur based on molecular weight prediction by means of native PAGE. In order to investigate the oligomeric assembly and the nature of interaction of subunits within the atypical structure, DpsADCT was subjected to X-ray crystallography as was carried out for DpsA and DpsC at Durham University. Conditions were much the same for DpsA and the final crystal structure diverged very little.

5.5.1 Structure validation of DpsADCT

The crystals of DpsADCT that were studied diffracted to 1.86 Å on beamline IO4 (Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxford, UK) and yielded one monomer per asymmetric unit with amino acids corresponding to 4-162 present in the tagged sequence which codes for a protein (including tag) of 168 amino acids. This means that 3 amino acids at the N-terminal region were not resolved in the crystal structure in addition to the hexa-His tag at the C-terminal. DpsADCT has no linker region between the end of the DpsADCT protein and the His tag. The protein sequence is given in Figure 5.8.

```
                                        A
MTHDLTPKYTVPGIEREAAGRLIGVLRLRLHALNDLHLTLKHVHWNVVGPH

              B                                  BC
FIAVHEMIDPQVDQVRDMADDVAERIAALGGVAQGTPGALVAERKWDDYS

                                  C                                  D
IGRADAIAHLGALDVVYTGVVEGMRAAVEEAGKIDPATEDLLIGQLRDLEQF

QWFVRAHLEHHHHHH
```

Figure 5.8: DpsADCT sequence annotated with tertiary structure. Highlighted in grey are the helical regions which are labelled above A through to D. Residues not resolved in the crystal structure are underlined with the His tag double underlined. Bold red residues are those that create the conserved ferroxidase centres.

The DpsADCT PDB 3D structure file was submitted to the Protein Structure and Validation Server version 4 (http://services.mbi.ucla.edu/SAVES/) and returned values of 99.334 % overall quality factor from the ERRAT program. 99.4 % of the residues were confirmed to be within the favoured region of a Ramachandran plot (created using http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) with just 0.6% of residues within the allowed region. These outputs are supplied in Appendices 23 and 24.

5.5.2 Overall Structure of DpsADCT

5.5.2.1 Monomeric Structure of DpsADCT – Preservation of the Four Helix Bundle

The DpsADCT monomeric structure is highly similar to DpsA with very little variation (Figure 5.9A). A four-helix bundle is retained indicating that the C-terminal tail does not influence tertiary structure formation. The helices labelled as A, B, BC, C and D are all present and composed of the same residues as DpsA as described in Chapter 4, furthermore the dimensions are equal to DpsA. The number of predicted salt bridges remains the same as DpsA using the same residues. However, one less hydrogen bond is predicted to strengthen the monomer, giving a total of 328 (Figure 5.9B)

5.5.2.2 Multimeric Structure of DpsADCT – An Unexpected Dodecamer

DpsADCT has an unexpected multimeric structure. Whilst *in vitro* evidence points to an assembly that is not comparable to a full Dps dodecamer within solution, the crystal structure yielded a complete dodecamer with 23-point symmetry (Figure 5.9 C+D). The three main structural elements of a dodecamer described in Chapter 4 were retained; the antiparallel dimers, the "ferritin like" trimers and the "Dps like" trimers that go towards forming the two types of pore. Whilst this assembly was unforeseen, it is not the first mutant Dps to display different oligomeric assemblies in solution and within crystals as the *M. smegmatis* mutant DpsΔC26 demonstrated trimeric assemblies in solution and decamers in the crystal structure (Roy et al., 2007). This points to an influence of the crystal environment and packing on protein structure (Eyal et al., 2005).

5.5.3 Structural Superposition of DpsADCT with DpsA and Structural Measurements

Superposition of proteins onto one another allows for the comparison of different conformations of protein folding and structure in order to visualise variations in structure. It also supplies numerical data (Root Mean Square Deviation-RMSD) that can be used to demonstrate the divergence between aligned amino acids. Since DpsADCT is a constructed mutant derivative of DpsA, DpsADCT was aligned with DpsA on a monomeric basis and also a dodecameric basis. The alignment at a monomeric level returned a RMSD of 0.385 Å between

158 atoms. Furthermore a RMSD of 0.387 Å at a dodecameric level highlights the conformational similarities between the two proteins (Figure 5.10C+D) yet still leaves a distance from which the two structures can diverge.

The external diameter of the DpsADCT cage measures 87.511 Å between Lys-96 of one subunit and Lys-96 of a subunit on the opposite side of the assembly. This compares with 87.202 Å measured on DpsA, a difference of 0.309 Å, which is a negligible value. Nevertheless a larger difference is found in pore size between the two structures. The N-terminal pore of DpsADCT measures 12.084 Å; this is 0.543 Å less than DpsA. There may be a structural reason behind this difference. However the significance this poses to the functional characteristics may be insignificant given the likelihood of different structural assemblies forming when in solution and during crystal packing.

5.6 Proteins, Interface, Structure and Assemblies Analysis (PISA)

Further comparisons between the characteristics of DpsADCT and DpsA can be made by analysing the effect, if any; the loss of the C-terminal tail has on the interfaces present within the DpsADCT assembly. The PDB file for DpsADCT was submitted to the PISA server for analysis and the results were then compared with DpsA. All the interfaces present in DpsA were detected; however there were slight variations between protein interface statistics.

Table 5.2 PISA statistics of DpsA and DpsADCT

	Interface type	Interface area, Å^2	$\Delta^i G$ kcal/M	$\Delta^i G$ P-value	N_{HB}	N_{SB}	CSS
DpsA	Dimer interface	1310.3	-7.9	0.552	12	5	0.660
	C-terminal trimer	512.0	-7.1	0.260	2	2	0.697
	N-terminal trimer	742.0	-3.5	0.552	8	7	0.295
DpsADCT	Dimer interface	1340.3	-7.7	0.554	12	2	1.000
	C-terminal trimer	447.2	-5.7	0.291	2	3	0.320
	N-terminal trimer	665.9	-2.0	0.628	9	6	0.311

Table 5.2 displays the average statistics for interface type of DpsA and DpsADCT. Before looking in depth at each interface it is possible to consider the Complexation Significance Score (CSS) which is an indication of how significant the formation of the interface is to the overall assembly. These numbers for DpsADCT are quite different to DpsA, where disruption of vital interactions originating from the C-tail regions shifts the significance away from the C-terminal trimer and onto the N-terminal trimer interface and dimer interface.

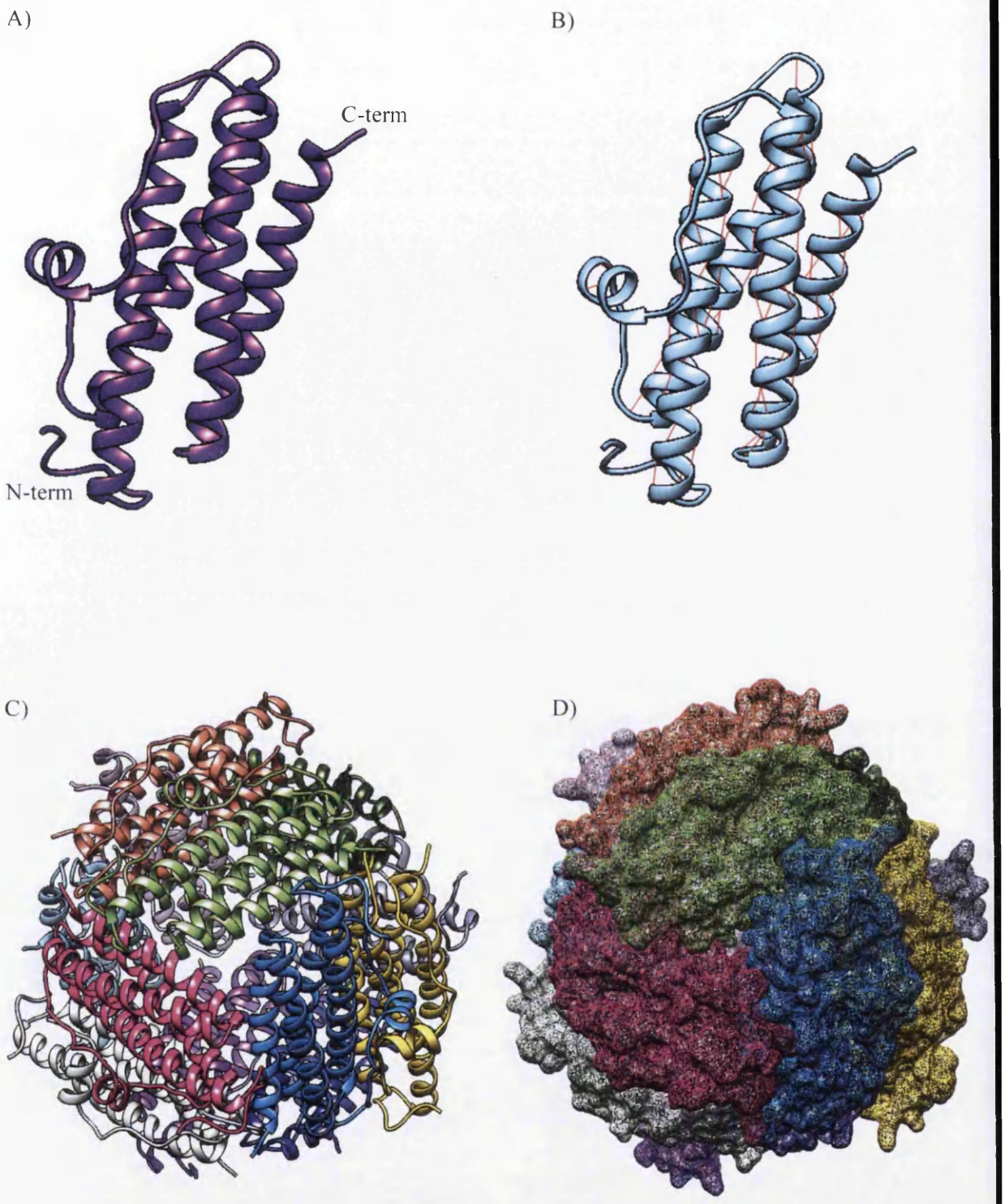


Figure 5.9: Crystal structure of DpsADCT at monomeric and multimeric states. Ribbon diagram of recombinant DpsADCT monomeric crystal structure highlighting the N and C-terminal tail extentions (A). The intra-monomer hydrogen bonds predicted to stabalise the four helix bundle (B). Crystal structure models of the DpsADCT dodecamer view from the N-terminal ferritin-like pore; ribbon (C) and mesh surface (D). Each subunit is tinted in a different colour highlighting the symmetry that exists in the dodecamer. Images were created using UCFS Chimera 1.7rc

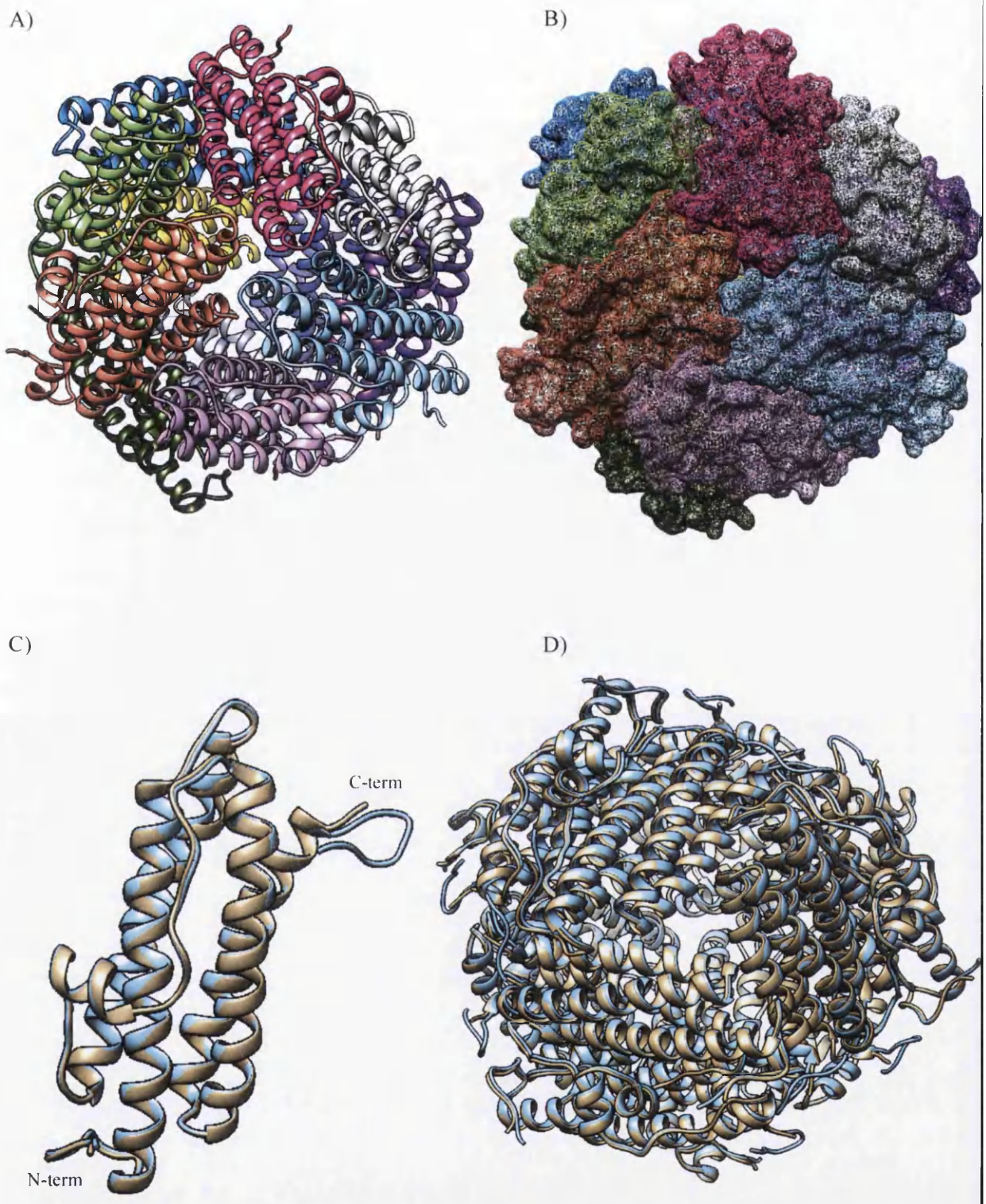


Figure 5.10: Crystal structure models of DpsADCT dodecamer and Structural Superposition alignment of DpsADCT and DpsA. View from the C-terminal "Dps-like" pore; ribbon (A) and mesh surface (B). Each subunit is tinted in a different colour highlighting the symmetry that exists in the dodecamer. Superposition alignments of DpsADCT (brown) onto DpsA (blue) at a monomeric (C) and dodecameric (D) level, RMSD are 0.385 Å and 0.387 Å respectively between 158 atoms. Images were created using UCFS Chimera 1.7rc

5.6.1 C-terminal Trimer Interface of DpsADCT is Less Significant to Oligomeric Assembly than it is for DpsA

The C-terminal trimer interface generates a CSS of 0.320, a score that is less than half that of DpsA (0.697). The interface itself in DpsADCT covers 64.8 $Å^2$ less interface area than the equivalent for DpsA and as a result the solvation free energy gain upon formation of the interface (Δ^iG) becomes more positive or, in other terms, slightly less hydrophobic. However the P-value statistic remains similar at below 0.500. This signifies the interface has a surprising "higher than average level of hydrophobicity" and signifies the interaction is specific. The hydrogen bonds within this region are maintained within the DpsADCT protein, specifically at the Glu-162 residue, the first residue of the C-terminal tail. H-bond and salt bridge formations between Glu-162 and Arg-104 exist in DpsADCT as they do in native DpsA. Nevertheless 3 interfacing amino acids at the C-terminal tail (Ala-164, Gly-165 and Gly-166) of DpsA are missing within DpsADCT. This also means that a single residue (Ala-105) present on the loop between the BC helix and the C helix of an adjacent subunit does not play a role in formation of the interface as it would in DpsA with the tail present. The structures with residues involved in creating the interfaces are given in Figure 5.11A+B with the raw output provided in Appendix 25.

5.6.2 N-terminal trimer interface of DpsADCT Becomes More Significant to Oligomeric Assembly

As described in Chapter 4, the C-terminal tail region of the protein is also involved in the formation and stabilisation of the N-terminal trimer (PISA Output Appendix 26). Of the 8 amino acids at the DpsA C-terminal tail that are visible within the crystal structure, 5 of these residues are interfacing residues, with 1 residue, Gly-165, involved in forming a hydrogen bond with Lys-8 of a symmetry related subunit. Of course Lys-8 is within the N-terminal tail of the symmetry related subunit. This hydrogen bond does not form in DpsADCT, it has no Gly-165; however, in its place is an H-bond interaction between Glu-162 and Tyr-9 that would otherwise not form in the full-length protein assembly. As seen with the C-terminal trimer interface, there are four residues in the C-terminal tail that would be involved in the formation of the interface (Gly- 165, Gly-166, Ala-167 and Leu-168) that are not present in DpsADCT. These cannot be involved in forming the N-terminal trimer interface, which is a reason why Lys-8 at the N-terminal is not involved in interfacing or H-bonding in DpsADCT. These regions would also affect the solvation free energy gain upon formation of the interface. The N-terminal trimer interface of DpsADCT has a more positive Δ^iG value; at - 2.0 kcal/M it could be considered to be less hydrophobic or have decreased positive protein affinity when compared to - 3.5 kcal/M of DpsA.

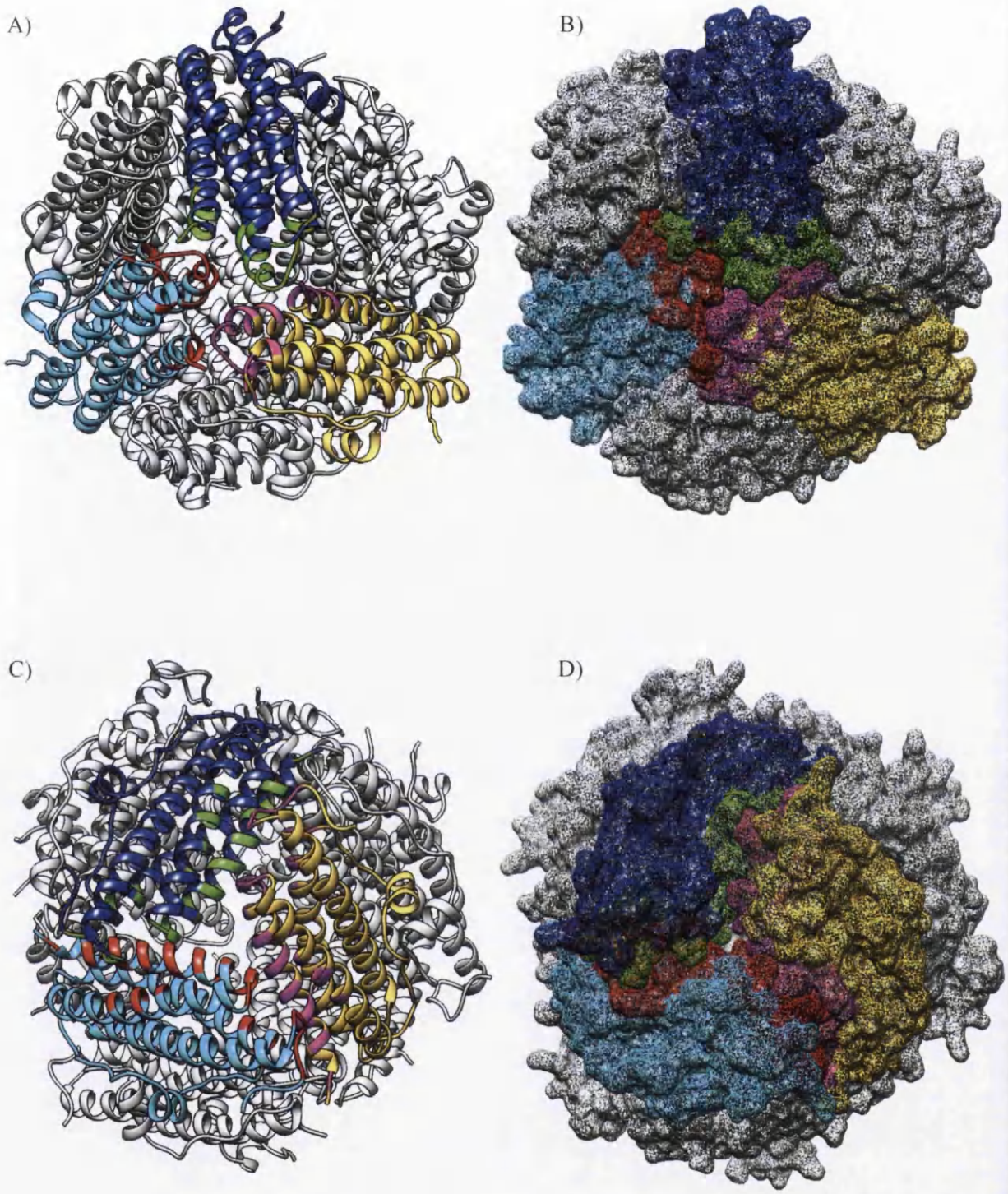


Figure 5.11: Structures of N and C terminal trimer interfaces of the DpsADCT dodecamer. N-terminal trimers form the "Ferritin-like pore" (A) and (B) and the C-terminal trimer which forms the "Dps-like pore" (C) and (D). 3-fold symmetry related monomers are represented by the colours blue, turquoise and gold, with the interfacing residues of the subunits coloured in green, red and pink. The other subunits are discoloured for emphasis. (A) and (C) are ribbon diagrams and (B) and (D) mesh surfaces. Images were created using UCFS Chimera 1.7rc

Loss of the C-terminal tail does not result in any more disorder in the N-terminal tail; evidenced by structure superposition analyses, DpsADCT N-tail adopts the same conformation as the N-tail in DpsA (Figure 5.10C). This is likely due to the preservation of interactions between the N-terminal tail and the helices within the bundle (Thr-10 H-bond with Arg-158 and Pro-12 H-bond with Arg-126) in addition to the Glu-162 to Tyr-9 interaction. Consequently the preservation of hydrophilic interactions compensates for the slight loss in hydrophobic interactions and so the CSS scores for this interface increase slightly making it more significant to oligomeric assembly than in DpsA.

5.6.3 Dimer interface of DpsADCT

The dimer interface is largely unchanged between DpsADCT and DpsA (Figure 5.12, Appendix 27). An increase of 30 $Å^2$ interface surface area suggests there may be slight neighbourhood movement at the A and B helices that may or may not be a direct consequence of the C-tail removal. This can also be visualised in the superposition of DpsA and DpsADCT (Figure 5.10), where residue Glu-57 has the highest RMSD value of the alignment. This residue along with the surrounding neighbourhood shows slight deviations in conformation and thus has increased distances from the residues in the reference structure of DpsA. Despite these small variations, the CSS score for the dimer interface of DpsADCT increases to 1.000. Since the C-terminal trimer has decreased in CSS an increase in dimer CSS highlights how imperative this interface is to dodecameric assembly in the DpsADCT mutant.

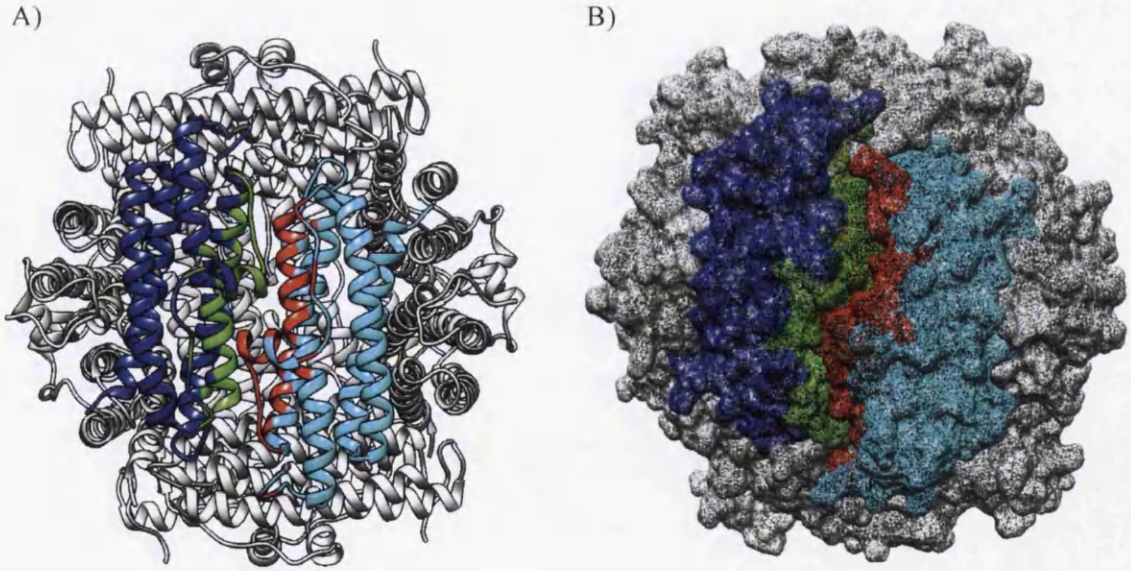


Figure 5.12: DpsADCT dimer interface. DpsADCT crystal structure models of dodecamers highlighting the dimer interface with anti-parallel subunits of a single dimer represented in differing colours (blue and turquoise). Highlighted in red and green are the opposing amino acids involved in forming the dimer interface. The other subunits are discoloured for emphasis. (A) Ribbon diagram and (B) mesh surface. Images were created using UCFS Chimera 1.7rc

5.7 Role of Dps tails maintaining assemblies

Chapter 4 documented how the PISA analysis allows the exact interfacing and interacting residues belonging to the tails to be explored. The PISA analysis of DpsADCT can allow the same exploration to be carried out paying particular attention to the residues from DpsA that interacted with the C-tail and how the loss of the C-tail impacts on these interfaces.

The residues of the DpsA N-tail were shown to not only interface and interact with the four-helix bundle of adjacent subunits but also with the C-terminal tail of an adjacent symmetry related subunit at the N-terminal trimer interface. The loss of the C-tail means that Lys-8 no longer contributes to the interface and no longer can hydrogen bond with Gly-165 as in DpsA. Whilst Tyr-9 also loses an interface with two residues (Gly-166 and Leu-168), a hydrogen bonding interaction occurs with Glu-162, which does not take place in DpsA (Figure 5.13). Whilst the N-terminal trimer interface loses some interface surface area due to the loss of C-tail, hydrophilic interactions persist. The interfacing residues are given in table 5.3 and this highlights how the loss of the C-terminal tail means the N-terminal tail interfaces with 5 less residues when compared with DpsA.

With Glu-162 being the only residue left of the C-terminal tail, its interactions with surrounding subunits as part of the interfaces were studied. As described previously the interface between Glu-162 and Tyr-9 still occurs and actually gains a hydrogen bond. This is a likely result of a slight change in conformations of Tyr-9 and Glu-162 in DpsADCT (Figure 5.13B+C).

Table 5.3: DpsADCT N-tail Residues Intersubunit Interactions

Given below are the N-tail residues and the residues that they interact with from a single symmetry related subunit per interface.

N-tail Residue	Dimer Interface	N-terminal trimer	Total Interfacing Residues
Asp-4	-	-	**0**
Leu-5	-	-	**0**
Thr-6	-	-	**0**
Pro-7	Arg-104	-	**1**
Lys-8	-	-	**0**
Tyr-9	-	Arg-158, Glu-162 (Hb)	**2**
Thr-10	-	Arg-158 (Hb)	**1**
Val-11	-	Glu-151	**1**
Pro-12	-	Thr-119, Val-122, Glu-123, Arg-126 (Hb), Gln-154,	**5**
Gln-13	-	Arg-126	**1**
Ile-14	-	Arg-126	**1**
Glu-15	-	-	**0**
		Total Interfacing residues	**12**

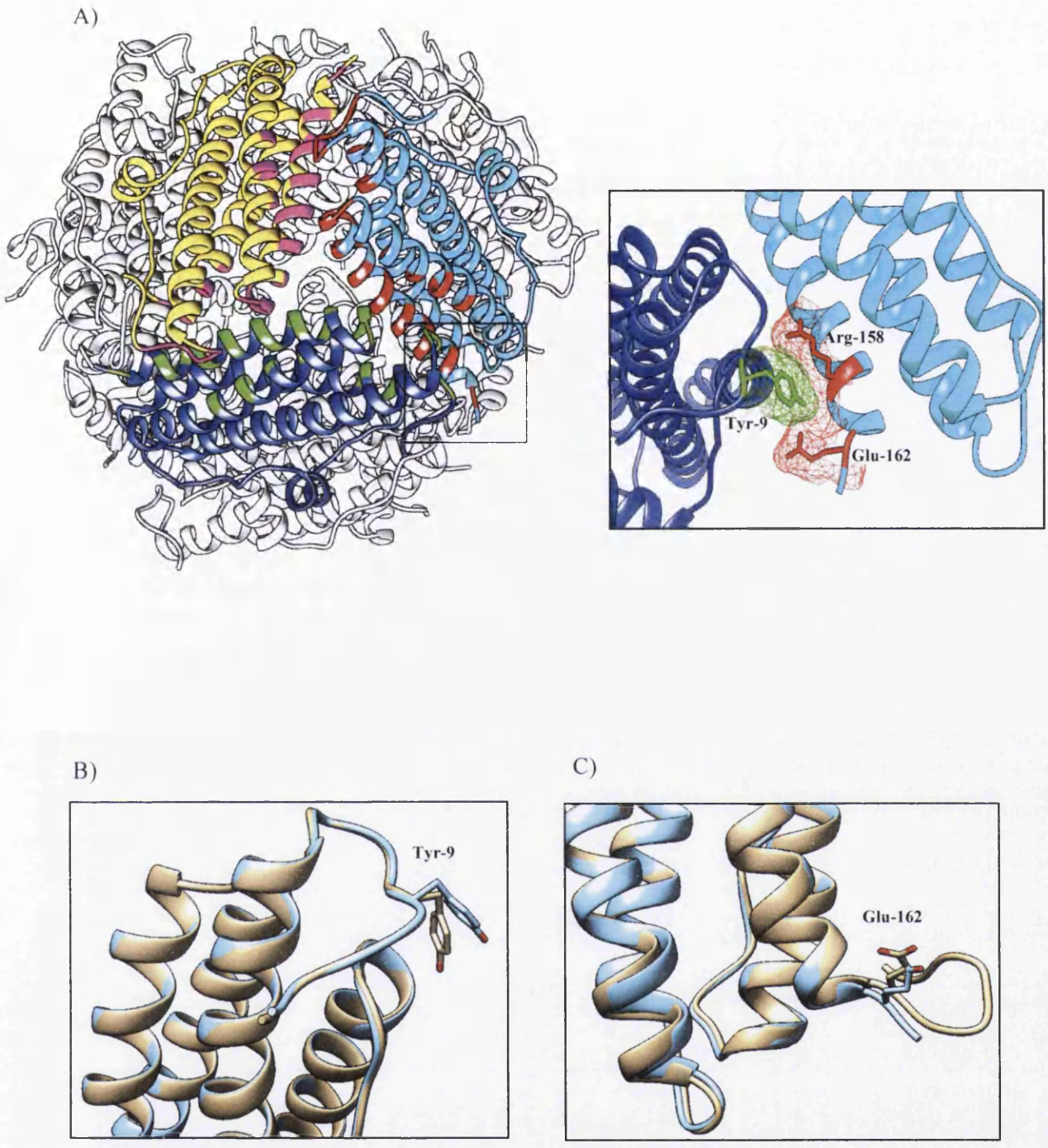


Figure 5.13: Important residues and structural alterations of DpsADCT tail regions. (A) DpsADCT interface between Tyr-9 from one subunit and Arg-158 and Glu-162 from a symmetry related subunit at the N-terminal trimer interface. A hydrogen bond also occurs between Tyr-9 and Glu-162 that does not happen in DpsA. (B) Conformation of Tyr-9 differs in DpsADCT (blue) when superimposed with DpsA (brown). (C) Glu-162 also shows conformation alterations in DpsADCT (blue) to DpsA (brown).

5.8 Summary of Protein Structure Manipulation and Analysis of Mutant Protein Structure results

Results in previous chapters pointed to the tails being influential in the ability of the Dps proteins to assemble into dodecamers. In order to elucidate the significance of each of the terminal tails to oligomeric assembly of the *Sc*Dps proteins, constructs were created in which the tail regions were removed. This chapter has assessed the detrimental impact that the loss of tails has had on assembly of the proteins and has explored how the abnormal assembly affects functional activity. Furthermore, the small BC helix of DpsA has also been removed and the impact on assembly assessed. The chapter has concluded with the elucidation and analysis of DpsADCT crystal structure and shown how it displays different oligomeric states in solution and in the crystal. A summary of the results from this chapter follows.

- Deletion of the tails from DpsA and DpsC severely affects oligomeric assembly
    - DpsADCT forms a predicted 8-mer in solution
    - DpsADNT forms a predicted 4-mer in solution
    - DpsADTM forms a predicted 3-mer in solution
    - DpsCDNT forms soluble aggregates that cannot be disrupted by urea
- DpsA deleted tail variants all maintain ferroxidase activity albeit with different progress rates.
- DpsA deleted tail variants preserve the ability to protect DNA against oxidative damage.
- Deletion of the BC helix disrupts proper dodecameric assembly, but multiple oligomeric assemblies of still occur.
- The crystal structure of DpsADCT displayed an unexpected dodecameric assembly with low RMSD values upon superposition alignment with DpsA at monomeric and multimeric levels.
- PISA analysis highlights some crucial residues, which may help to maintain the structure, but highlights a decrease in complexation significance score for the C-terminal trimer indicating the importance of the C-tail to this structure.

Chapter 6

Exploring The Nanotechnological Applications Of DpsA

6.1 Introduction

The fabrication of novel biologically derived nanomaterials is attracting great interest from within the science community. As a well-studied protein, ferritin is a prominent focus for the generation of nanoscale materials particularly because of its ability to mineralise metals, encapsulating them within its internal cavity (Okuda et al., 2003). Given that the proteins have the ability to produce paramagnetic nanoparticles in a size constrained environment, they have been exploited for the development of a range of applications from biomedicine (Doll et al., 2013) to electronics (Ko et al., 2011). Despite there being an effort to explore the potential of ferritin, very little attention has been given to Dps even with the protein family possessing interesting characteristics such as smaller diameters and high temperature stabilities (Franceschini et al., 2006).

The crystal structures of DpsA and DpsC highlighted interesting features at the nano scale of these spherical protein cages. Possessing external diameters of roughly 10 nm and internal cavities that measure around 4 nm, these dimensions lie well within the threshold of what defines nanotechnology. Furthermore, the ability to deposit an iron oxide core increases their potential for use as nanotechnological agents. The differences in stabilities between DpsA and DpsC also evoke another avenue to explore, as stability in a range of environments both physiological and extreme would expand the list of possible applications.

This chapter explores the potential for Dps to be developed as biotechnological tools, specifically in a biomedical context. The ability to control and direct how the proteins assemble is investigated and the potential for functionalising the protein cages is explored.

6.2 Hetero-oligomeric assembly

The removal of N – and C – terminal tails from the four-helix bundle of the DpsA and DpsC proteins resulted in the incapacity of the proteins to assembly into dodecamers. Whilst this observation has also been described for the Dps proteins of M. *smegmatis* (Roy et al., 2007) and *D. radiodurans* (Bhattacharyya and Grove, 2007), very little has since been reported into how this phenotype can be harnessed for use in biotechnology. The self-assembling nature of Dps is a high hurdle when it comes to manipulate how, when and where these proteins assemble *in vitro*, particularly for the highly stable assembly of DpsC that subsequent to assembly resists dissociation. The non-assembling tail-less variants of DpsA provide an interesting opportunity to investigate just how easy it would be to gain an assembly modulating technique. Furthermore, it remains to be seen if the full-length proteins are capable of re-establishing

dodecameric assembly of the tail less variants by creating chimeric dodecamers. This also raises the concept of hetero-oligomeric assemblies forming between different proteins *in vivo*; DpsA and DpsB in the same dodecamer, for example. This would serve as an intricate means of modulating the function of the proteins *in vivo* and would explain why there has been no evidence for DpsB assembling into a dodecamer.

6.2.1 Disrupting and re-establishing of the Dps oligomeric state

In order to explore if the incorporation of tail-less variants into dodecamers using the full length protein was achievable, a means of disrupting the dodecamers was first required. The research in previous chapters provided useful data on deciding how this would be approached. DpsC was found to be highly stable in a range of conditions and was resistant to oligomeric disruption by urea. The DpsCDNT mutant was shown to form large soluble aggregates incapable of entering a native PAGE gel and was also resistant to denaturation by urea and so these proteins were not considered for this investigation. The assembly of DpsA, on the other hand, was easily disrupted and dissociated by incubation with urea at a minimum concentration of 6 M. The first phase of this experiment was to explore the denaturing dissociation and subsequent renaturing and re-association of the DpsA dodecamer. After incubation with 8 M urea for 24 hours to allow complete dissociation, the urea was removed by gel filtration. The eluates were collected in four fractions to find if there was any difference in refolding between the initial part of the sample injection to flow through and the last to flow through. Figure 6.1A+B display the results of the refolding experiment where native PAGE was carried out immediately after buffer exchange and 24 hours post buffer exchange. Equal quantities of protein were loaded in each lane. The gels demonstrate how DpsA's dodecameric assembly can be disrupted by 8 M urea and after the removal of the urea the protein can reassemble. There was no difference between the elution fractions. However there was a difference in relative quantity of protein reassembled between the times. After 24 hours, a stronger dodecameric band can be visualised with slightly less intermediate species present. This demonstrates how hetero-oligomeric experiments would require a time scale of at least 24 hours to allow full reassembling of the proteins to occur.

DpsADCT showed a non-dodecameric assembly in solution. Assessment of DpsADCT's assembly after the removal of urea was carried out and proved that whilst it has the ability to reassemble it cannot reassemble into a dodecamer. Figure 6.1C displays how the assembly can be disrupted and reassembled.

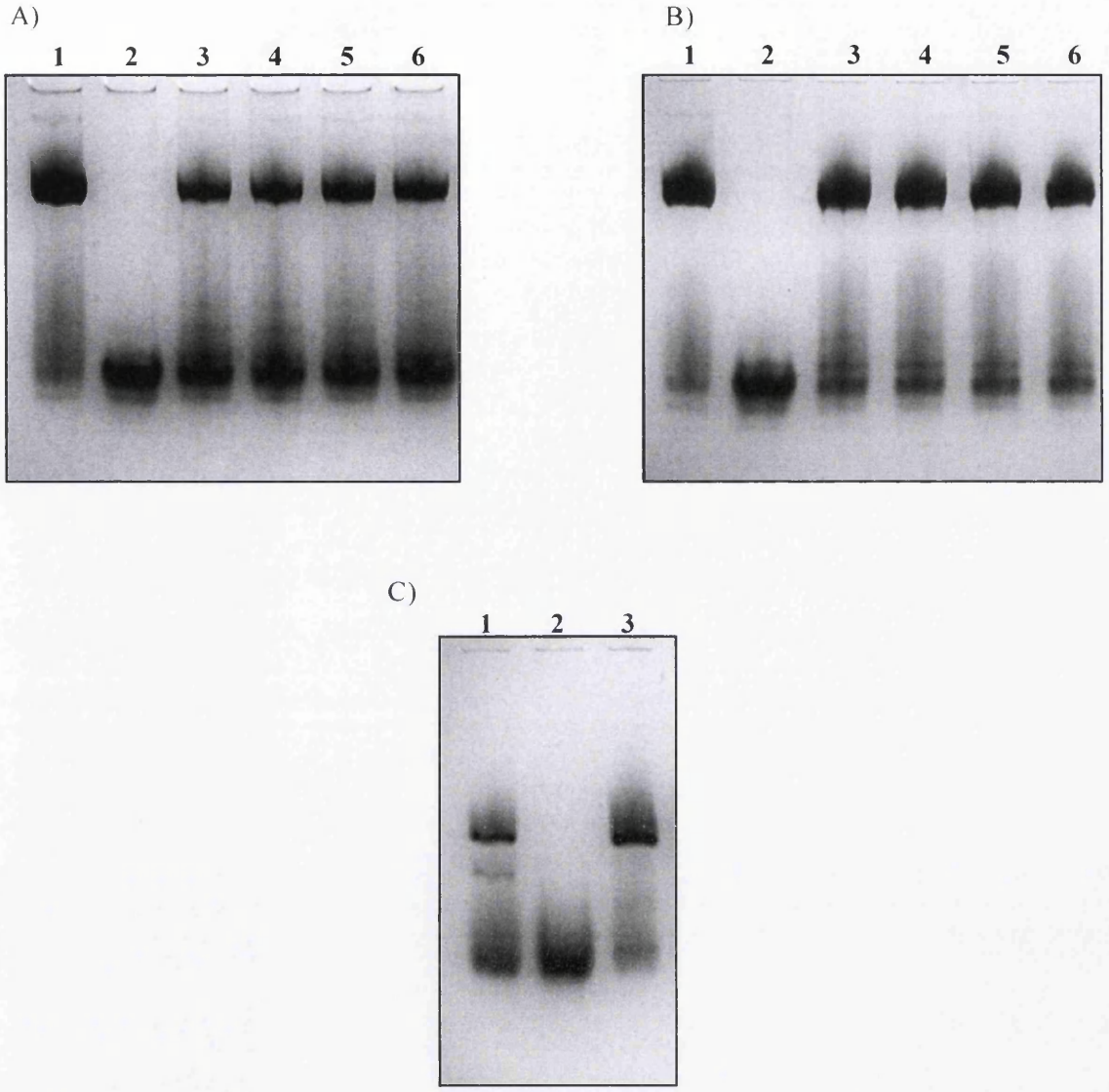


Figure 6.1: Urea mediated dissociation and re-association of the DpsA dodecamer. Immediately after (A) and 24 hours post urea removal (B) of urea DpsA can be refolded. Lane 1 – DpsA, lane 2 – 8 M Urea DpsA, lanes 3 to 6 – refolded DpsA fractions in order of flow-through. (C) DpsADCT urea mediated dissociation and re-association: Lane 1 – DpsADCT, lane 2 – DpsADCT 8M Urea and lane 3 – DpsADCT refolded.

6.2.2 Hetero-oligomeric assembly of DpsA and DpsADCT

Having demonstrated that both DpsA and DpsADCT can have their oligomeric assemblies disrupted and reformed, hetero-oligomeric assembly was investigated. To ascertain if hetero assemblies between DpsA and DpsADCT were formed, samples of DpsA and DpsADCT in 8 M urea were mixed together in equal proportions and the urea removed as described. Following 24-hour incubation a 2-dimentional approach was employed as described in Experimental Procedures (Section **2.11.5**).

The first dimension native PAGE reveals the appearance of multiple species, which migrate in accordance to a dodecameric assembly (Figure 6.2A). Within the mixture lane, there is very little evidence of DpsADCT refolding as it does in the absence of DpsA which points to its inclusion in the higher oligomeric bands. The multiple dodecameric bands may represent different ratios of DpsADCT to DpsA, but could also result from incomplete refolding or assembly. Extracting the bands from a replicated lane (Figure 6.2B) and also extracting an entire lane (Figure 6.2C) gave positive results in the second dimension of SDS PAGE. The dodecameric bands from the native PAGE gel presented two clear bands on the SDS PAGE gels with masses appropriate for the monomers of each protein. This suggests hetero-oligomeric assembly between DpsA and DpsADCT does occur but the proportions of DpsA to DpsADCT may be random. It also suggests the loss of the C-tail that causes loss of dodecameric assembly can be compensated by the inclusion of full-length protein monomers. The complexing of the DpsA (with its C-terminal) tails to DpsADCT can provide some of the interactions that occur at N-terminal trimer interface and the C-terminal trimer interface as discussed in Chapter 4, estabilising the chimeric dodecamer.

6.2.3 Hetero-oligomeric assembly of DpsA and DpsADTM

Hetero-oligomeric assembly between DpsA and DpsADTM was also tested and yielded positive results. Figure 6.3A demonstrates that the non-assembling DpsADTM does not show any loss of oligomeric state once placed in 8 M urea; equally it does not possess the ability to form higher state oligomeric assemblies once the 8 M urea is removed. However upon the removal of urea from the DpsADTM-DpsA mix, the first dimension displayed another set of high oligomeric bands with a decrease in intensity of the low band corresponding to DpsADTM. Processing the high oligomeric bands into the second dimension revealed two bands with masses corresponding to DpsA and DpsADTM (Figure 6.3B). The ability to take a non-multimeric Dps and incorporate it into a dodecameric assembly provides a means of being able to control the protein's self-assembling nature. It also suggests that the presences of some tails within the dodecameric assembly are sufficient to strengthen the interfaces found at the N – terminal and C – terminal interfaces.

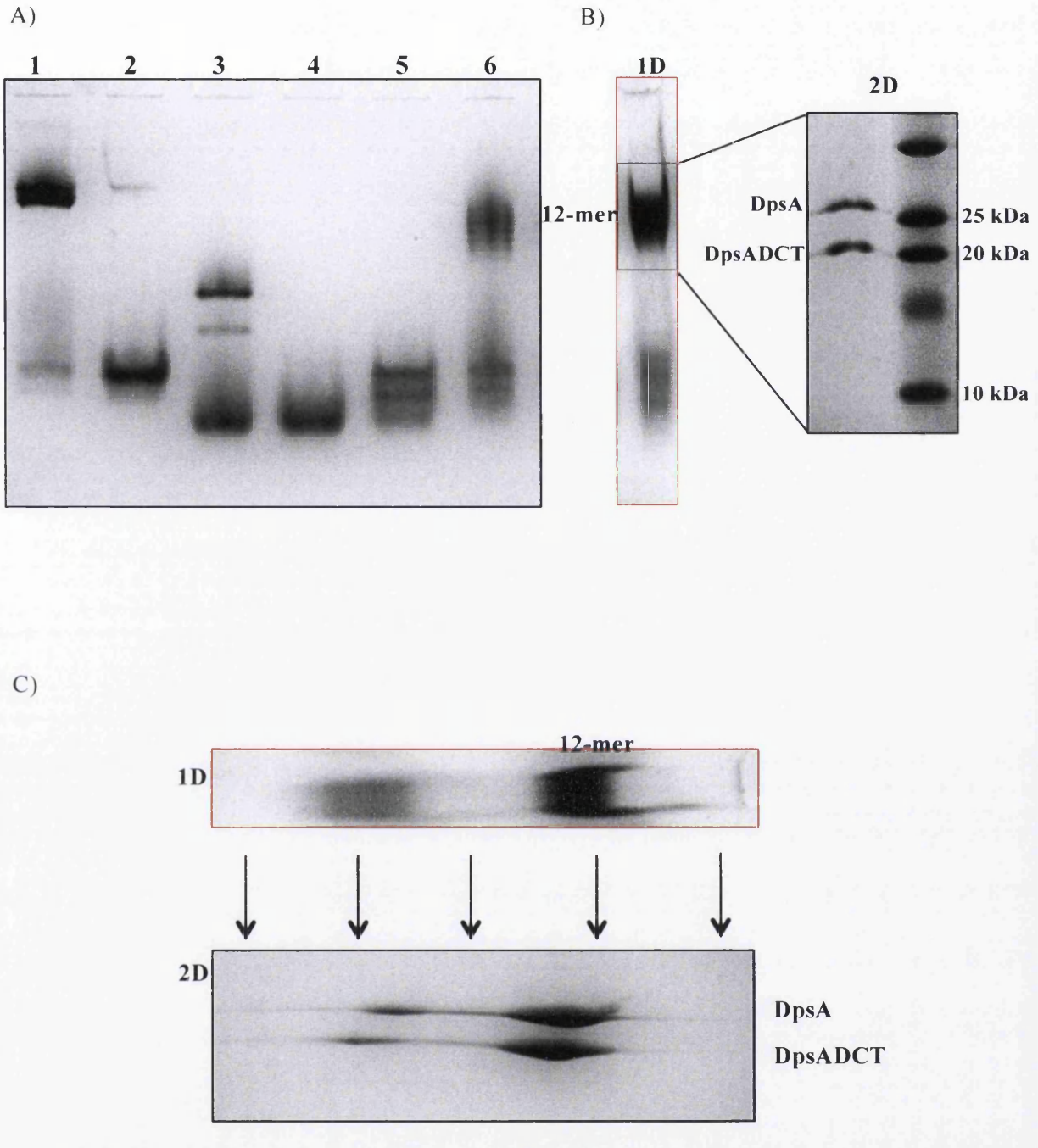


Figure 6.2: Hetero-oligomeric assembly of DpsA and DpsADCT. (A) 1st Dimension native PAGE, lane 1 – DpsA, lane 2 – DpsA 8 M urea, lane 3 – DpsADCT, lane 4 – DpsADCT 8 M urea, lane 5 – DpsA and DpsADCT mix 8 M urea and lane 6 – DpsA and DpsADCT refolded. (B) Replica lane of 1st dimension (red boxed, 1D) loading an increased amount of refolded DpsA and DpsADCT with black box indicating the extracted area processed for 2nd dimension (2D-SDS PAGE) using Biorad Precision PlusProtein™ Kaleidoscope as a standard. (C) 1st and 2nd dimensions using entire native PAGE lane rotated 90 ° and placed on top of an SDS PAGE gel.

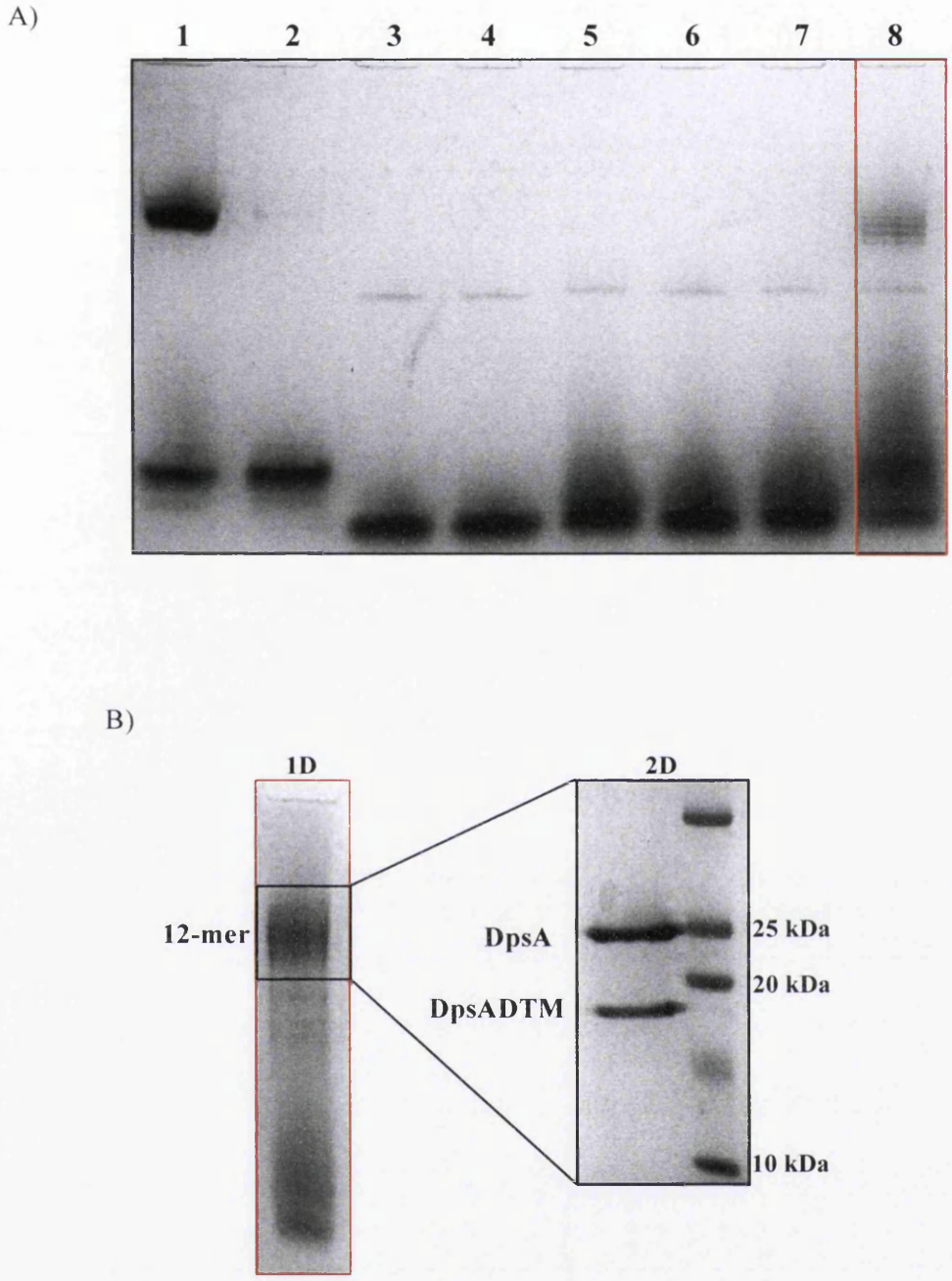


Figure 6.3: Hetero-oligomeric assembly of DpsA and DpsADTM. (A) 1st Dimension native PAGE, lane 1 – DpsA, lane 2 – DpsA 8 M urea, lane 3 – DpsADTM, lane 4 – DpsADTM 8 M urea, lane 5 – DpsADTM refolded 1, lane 6 – DpsADTM refolded 2, lane 7 – DpsADTM refolded 3 and lane 8 – DpsA and DpsADTM refolded. DpsADTM shows no sign of high multimeric assembly on its own, but is capable of being incorporated into a dodecamer with DpsA (B) Replica lane of 1st dimension (red boxed, 1D) with black box indicating the extracted area processed for 2nd dimension (2D-SDS PAGE) Biorad Precision PlusProtein™ Kaleidoscope standard was used.

6.2.4 Hetero-oligomeric assembly of DpsA and DpsB

DpsB was found to be incapable of assembling into a dodecamer under the conditions tested both *in vivo* and *in vitro*. The results from the last two sections have demonstrated that a dodecamer can be formed from a mixture of tailed and tail-less species, taking non-assembling proteins and incorporating them into the multimeric complex. DpsB has short tails and in some way the DpsADTM oligomeric behaviour mirrors what is observed for DpsB. For that reason the ability to form hetero-oligomeric assemblies with DpsA was explored. The ability of DpsB to assemble into higher order complexes after removal of urea was also tested. DpsB cannot form multimeric assemblies after the removal of urea, as evidenced by a lack of band shift in the DpsB refolded lane Figure 6.4A.. The re-assembling of DpsA from urea was also tested as a further control. Since DpsA reassembly occurs, it also serves to validate that the experimental set up works (Figure 6.4A).

The removal of urea from a DpsA and DpsB mixture generated a multimeric assembly that did not provide the "ladder" feature previously seen. However there was also the presence of strong lower bands on the gel, which points to non-assembling proteins (Figure 6.4A). The second dimension proved that the dodecameric assembly consisted of just a single DpsA species; no DpsB could be detected associated with DpsA. Conversely the lower bands present in the 1st dimension gel can be attributed to DpsA and DpsB (Figure 6.4B). This result suggests that hetero-oligomeric assembly between DpsA and DpsB is not likely to occur and the concept that it may act as an activity modulating mechanism *in vivo* is improbable. On the other hand this experiment does act as a negative control experiment to the other hetero-oligomeric assemblies of DpsA with DpsADTM and also with DpsADCT. The result also points to the importance of the four-helix bundles for the formation of the multimeric assembly. As much as the tails have been shown to be important to oligomeric assembly, this experiment has shown that there is still a need for the four-helix bundle interactions to occur in order to create the dodecamers. Since DpsB has a different amino acid sequence it makes it impossible to form interactions across the three types of interfaces described in the PISA analysis of Chapter 4. It is these structures that go toward holding together the dodecamer and as a consequence the presence of DpsA's tails has little effect on the ability for DpsA and DpsB to co-assemble. DpsADCT and DpsADTM share the same four-helix bundles, accordingly the trimers and dimers are capable of forming with DpsA and the tails brace the structure, holding it together.

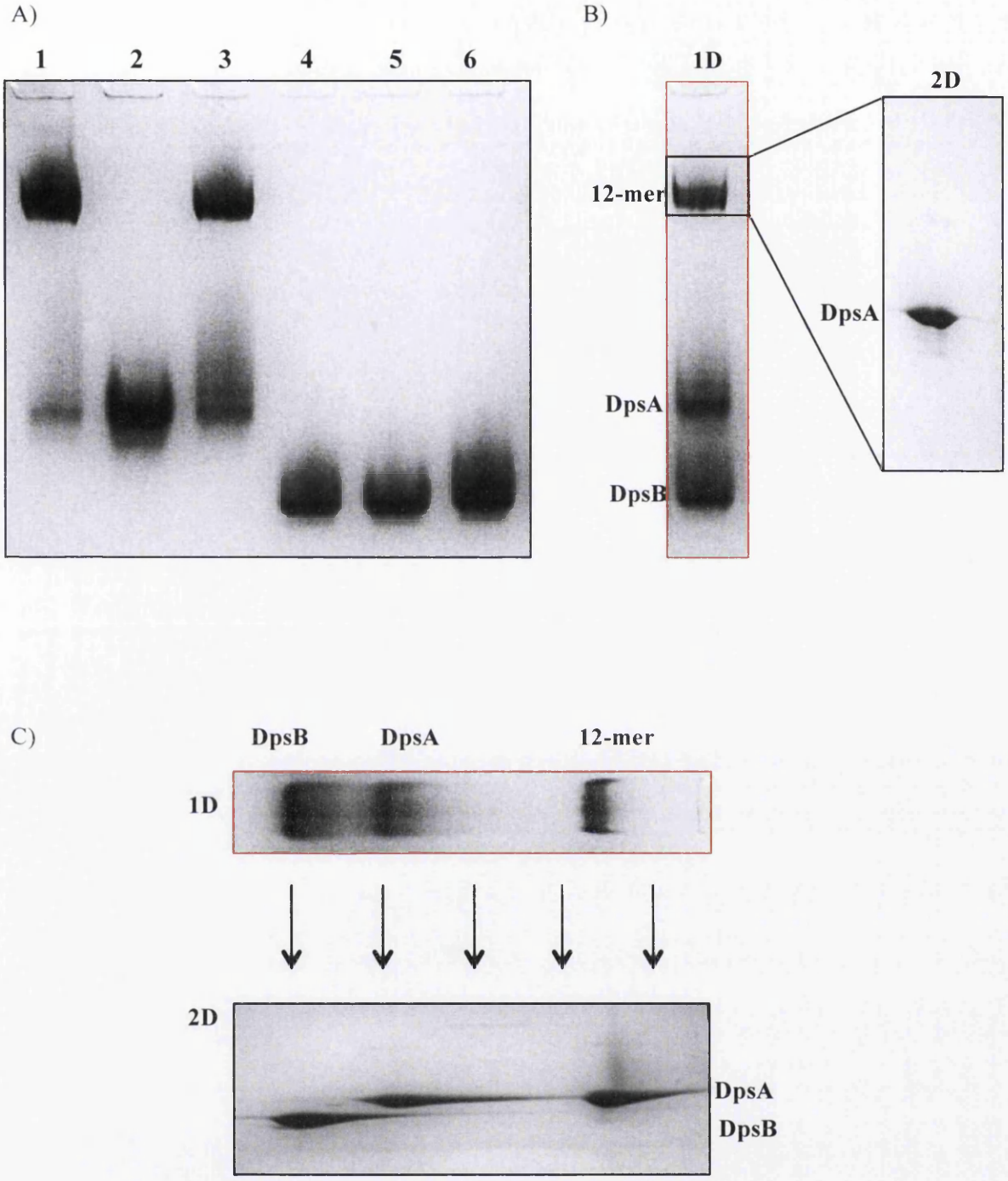


Figure 6.4: Hetero-oligomeric assembly of DpsA with DpsB. (A) 1st Dimension native PAGE, lane 1 – DpsA, lane 2 – DpsA 8 M urea, lane 3 – DpsA refolded, lane 4 – DpsB, lane 5 – DpsB 8 M urea and lane 6 – DpsB refolded. (B) 1st Dimension native PAGE DpsA and DpsB refolded (red boxed, 1D) with black box indicating the extracted area processed for 2nd dimension (2D-SDS PAGE) and the appearance of a single band (DpsA). (C) 1st and 2nd dimensions using entire native PAGE lane rotated 90 ° and placed on top of an SDS PAGE gel.

6.3 Functionalising the DpsA cage

Exploiting the protein cages to have biomedical uses often relies on the addition of functional groups to the protein frame at three distinct regions; the exterior, the interior and inter-subunit surfaces. The functional groups themselves can be extremely diverse and can be of organic or inorganic origin. Purposes such as MRI contrast agents, specific cell targeting and encapsulation of drugs have all been documented within the literature as being potential nanocage related technologies (Flenniken et al., 2009). In order to apply the knowledge gained from the research into the biochemical and structural properties of the proteins, the investigation into functionalising the Dps nanocage was carried out using DpsA as a platform. A series of functional groups were chosen as prototypes to test the technology in the effort to explore the feasibility of chimeric multi-functionalised Dps paramagnetic nanoparticles.

6.3.1 Dps-mCherry translational fusion

The red fluorescent protein mCherry has been used throughout molecular biology as a quantitative reporter often used to monitor the expression and localisation of proteins *in vivo* (Pradel *et al.*, 2007), (Facey et al., 2009). As a multimeric protein, it's development into a fluorescent monomer by directed mutagenesis improved its ability to be used as a translational fusion, minimising problems with target protein function (Shaner et al., 2004). The mCherry protein has already be used to monitor the localisation of DpsA expression *in vivo* without any obvious phenotypic problems at a protein or cellular level (Facey et al., 2011). However its oligomeric assembly was not explored. Therefore the reasons for its use as a prototype to functionalise DpsA were twofold. Primarily the experimental data will demonstrate if it is possible to functionalise the external surface of the DpsA cage without encountering major structural and function problems that may affect cage assembly and ferroxidase activity. Secondly it provides a meaningful tool that has practical use as a biotechnological application if used in conjunction with other functional groups that possess biological recognition and binding properties. This can then be applied in various ways, for example conferring a fluorescent signal to the surface of a specific cell.

In order to functionalise DpsA, a C-terminal translational fusion was used. Previous results showed that the N-terminal region couldn't be manipulated; furthermore the crystal structures show the C-tail protrudes into the solvent. The upshot of the C-tail conformation is the "pushing" of the fusion away from the interfacing regions of the protein. A plasmid, pDpsAmCh, which coded for the translational fusion construct, was used in the experiment. The plasmid was created by amplifying the mCherry coding sequence by PCR using primers which contained relevant restriction enzyme recognition sites for downstream cloning; the forward primer containing a *Sac*I recognition sequence and the reverse primer containing a *Not*I

recognition sequence and removing the native stop coding. The PCR product was cleaned and ligated into pDpsA14 prepared with a *Sac*I/*Not*I digest. This creates DpsA with mCherry fused at the C-terminal followed by a hexa-histidine tag. The construct encodes a monomeric protein with a total mass of 48.69 kDa of which mCherry contributes 26.94 kDa. Presuming that dodecameric assembly would still occur, the predicted mass of the dodecamer would be 584.28 kDa. The protein has a predicted iso-electric point of 5.5 only slightly higher than DpsA alone that is 5.41 when tagged. Despite the placement of the fusion to the C-terminal, eventually placing it on the external surface of the dodecamer, there is still a real possibility of steric hindrance preventing assembly of the Dps dodecamer.

The pDpsAmCh plasmid (Figure 6.5A), was transformed into BL-21 (DE3) cells for protein expression. Protein purification was achieved from the soluble fraction and displayed a strong pink colour within the solution. The protein's assembly was analysed by native PAGE. In accordance with the migration of DpsA in a native gel being persistently larger than its predicted molecular weight, it was expected that DpsAmCh would behave in the same manner. Figure 6.5B demonstrates that DpsAmCh migrated above the 720 kDa Apo ferritin band 1 of the Invitrogen NativeMark™ molecular weight standard.

The ability to deposit an iron oxide core within the DpsAmCh complex would reveal two key factors about the protein; that ferroxidase activity is maintained showing that at least the antiparallel dimers of DpsA are still capable of forming, and also that the protein must have some form of cavity in which to mineralise and deposit the ferric iron. Feeding the protein solution with ferrous iron tested this, processing the protein as previously described. The iron loaded DpsAmCh together with an unloaded control were then resolved by native PAGE. Figure 6.5C demonstrates the range of ways the gel was visualised and imaged. The high concentration of protein required to give a signal in the iron staining process meant that the concentration of mCherry within the gel was enough to give strong visible pink bands. Furthermore, both bands from the apo and holo DpsAmCh proteins fluoresced when subjected to an excitation wavelength of 587 nm (emission 610 nm). Additionally, a positive band from the iron staining process proved that not only is ferroxidase activity maintained but a cavity to deposit the iron oxide also exists within the assembly.

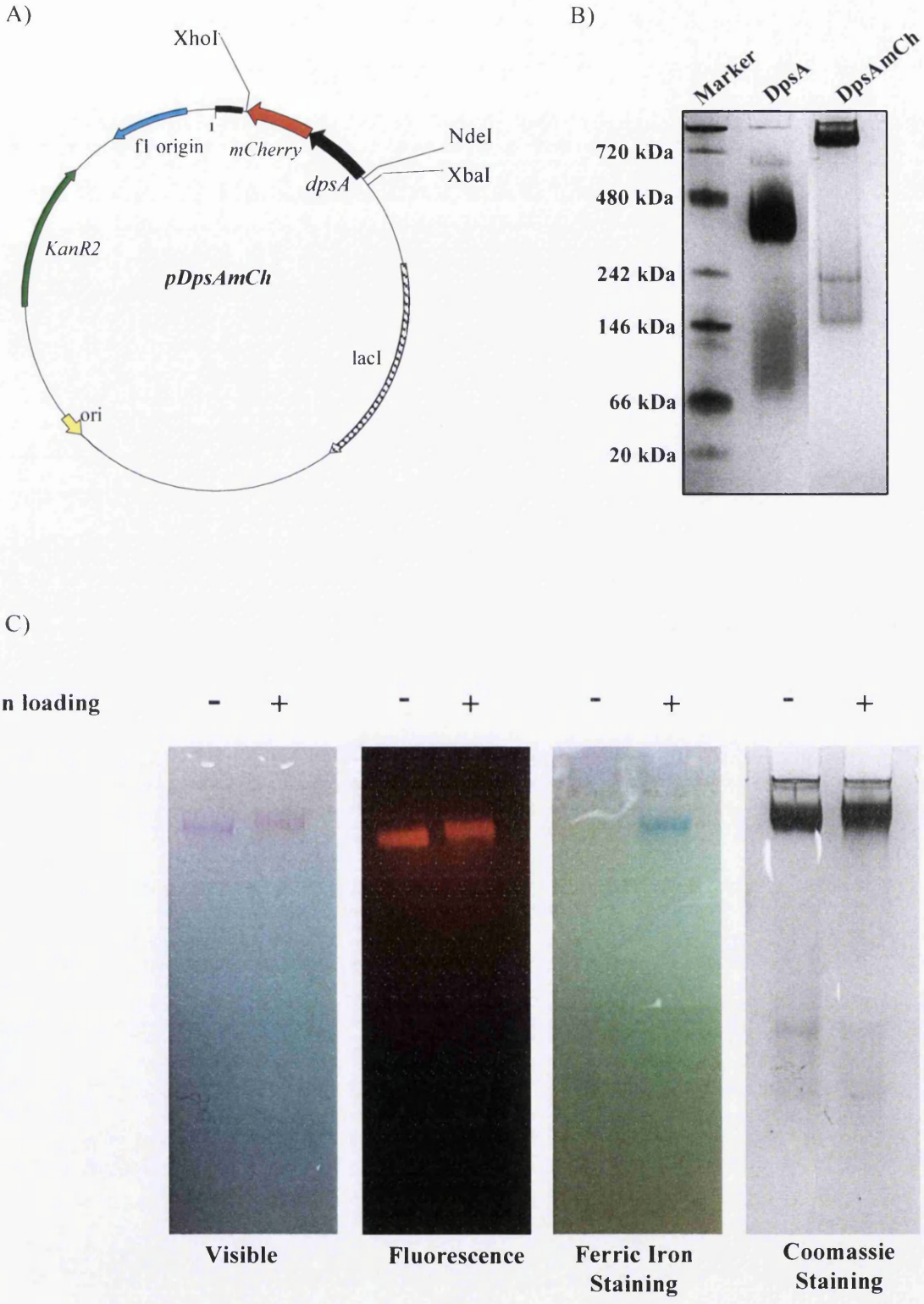


Figure 6.5: Assembly and functional analysis of DpsAmCh functionalised cage. The plasmid map of pDpsAmCh shows the C-terminal translational fusion from DpsA to mCherry followed by hexa-histidine tag (A). Map not drawn to scale. (B) Composite image of DpsAmCh resolved by native PAGE and stained with coomassie using the Invitrogen NativeMark™ standard. (C) DpsAmCh iron loaded and control resolved by native PAGE and visualised in four different ways; visible pink bands, fluorescence, ferric iron staining and coomassie staining.

6.3.2 Hetero-oligomeric assembly of DpsA and DpsAmCh

The hetero-oligomeric assembly of DpsA and DpsAmCh was tested to find if the addition of the functional fusion at the C-terminal would affect the ability to create chimeric dodecamers. Hetero-oligomeric assembly has been shown to occur between non-dodecameric species and full length DpsA using the tails to compensate for an otherwise unfeasible assembly. However DpsAmCh already does assemble, therefore the ability to disrupt the DpsAmCh assembly using 8 M urea was initially explored. A gradient of urea concentrations were tested using 2, 4, 6 and 8 M concentrations.

The results show how only 4 M urea is required to dissociate the complex (Figure 6.6A). This is less than required for DpsA dissociation and suggests that whilst the DpsAmCh complex is likely to form a dodecamer, the interactions present at the interfaces of the DpsA structure may be weakened by the addition of the mCherry fusion. When dissociated by 4 M urea, DpsAmCh migrates with a mass just above the 146 kDa lactate dehydrogenase band of the molecular weight standard. This suggests that the urea takes the assembly down to a trimeric state. However an increase in urea concentration to 8 M urea does not decrease oligomeric state any further.

The hetero-oligomeric assembly analysis of DpsA and DpsAmCh showed that chimeric assemblies were created composed of DpsA and DpsAmCh within the same complex. Very little DpsAmCh was refolded as a homo-oligomeric DpsAmCh species. However, a distinct ladder effect with masses indicative of multimeric assembly points toward the creation of chimeric assemblies. Pink bands were barely visible, and given the lower protein concentration this was to be expected. The gap between the different multimeric bands reflects the difference in monomeric molecular weight between DpsA and DpsAmCh (22.37 kDa vs. 48.69 kDa). The second dimension actually retains the boundaries between the bands. This not only confirms that DpsA and DpsAmCh were likely to be interacting, but also gives the resolution to show that DpsA also formed a homo-dodecamer. It also signals that the lowest molecular weight chimeric assembly contained more DpsA than DpsAmCh, which is to be expected. As the assemblies increased in molecular weight, the abundance of DpsA was slightly decreased. This also substantiates the notion that different ratios of DpsA and DpsAmCh occur within the dodecamers. On the whole this proof of concept was successful in proving that DpsA can be functionalised using a C-terminal translational fusion approach. It also shows how the hetero-oligomeric assembly may work as means of controlling the number of functional groups per dodecamer, or in creating multi-functional cages that contain more than one type of functional group on the surface of the dodecamer.

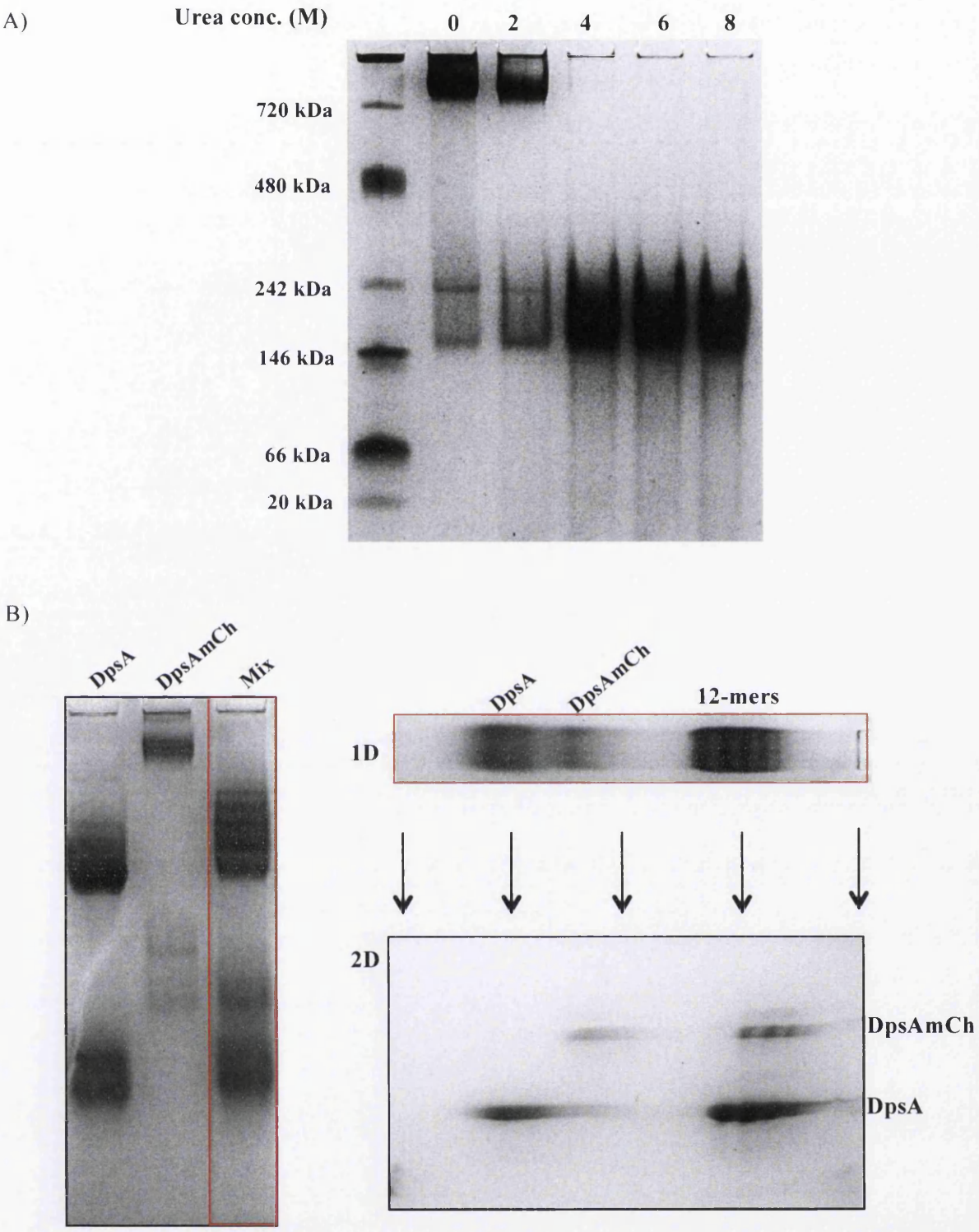


Figure 6.6: DpsAmCh sensitivity to urea and Hetero-oligomeric assembly of DpsA and DpsAmCh (A) DpsAmCh sensitivity to urea resolved by native PAGE and coomassie stained; 4 M urea is enough to disrupt full assembly. Urea concentrations are given above the lane. Invitrogen NativeMark™ was used as a standard. (B) Hetero-oligomeric assembly of DpsA and DpsAmCh is demonstrated with a two dimensional approach. 1st Dimension native PAGE (red box – 1D) shows a "ladder like" assembly. The 2nd dimension (2D – SDS PAGE) where the 1D lane was rotated 90° and placed on top of an SDS PAGE gel shows two proteins present within the assembly.

6.4 Functionalising DpsA with lysins

Functionalising the DpsA protein cage with the fluorescent mCherry protein showed that translational fusions could be an option for creating functionalised and assembled protein nano-cages that retain biological activity in the form of ferroxidation and iron binding ability. Furthermore, the function of the fusion, in this case fluorescence, was also maintained. However this fusion does not have a biomedical purpose as it is. In order to transform this technology into one that has biomedical significance, consideration has to be given to the nature of the functional group applied. The choice of functional group heavily depends on the purpose and application of the technology. As previously described, there are a number of potential uses for functionalised nanoparticles, however in this case the focus will be based on selective detection, targeting and killing of pathogenic bacteria.

Phage peptidoglycan hydrolases (lysins) are enzymes used by dsDNA bacteriophages to lyse the cells they infect as a final stage of their multiplication cycle (Matsuzaki et al., 2005). Their potential as antimicrobial agents have been recognised for decades with the use of bacteriophages predating the discovery of antibiotics (Sulakvelidze et al., 2001). Yet because of the effectiveness of antibiotics, the study of phages and their lytic methods was virtually abandoned. However overuse of antibiotics has subsequently resulted in the generation of widespread antibiotic resistance (Davies and Davies, 2010). For this reason, research interest has returned to bacteriophages and specifically endolysins in the search for therapeutic control of pathogenic bacteria.

The enzymatic effect of lysins is achieved by digesting one of the four bonds within peptidoglycan. It is this mode of action that renders this family of proteins effective against gram-positive bacteria (Fischetti, 2010). The specificity shown by endolysins owes itself to the domain structure of the proteins. Generally divided into two domains; the N-terminal domain has catalytic hydrolase activity while the C-terminal domain is involved in cell recognition and binding, hence is known as the cell binding domain (CBD) (Loessner, 2005). Interestingly, the CBD of a lysin is not always required for the enzymatic activity of the N-terminal domain (Low et al., 2005). On the other hand the CBD has been demonstrated to retain specific recognition and binding ability of cells in the absence of the N-terminal domain (Kretzer et al., 2007). The current interest and generation of both biochemical and structural data pertaining to endolysins provides an opportunity to harness the knowledge and apply it to Dps nanotechnology where the nano-cage module of the fusion would not only contribute its ability to generate a paramagnetic core for magneto detection/targeting purposes, but also would define a scaffold for the coupling of multiple functional groups which enhance and/or diversify the utility of the functionalised fusion.

6.4.1 *Clostridium difficile* Targeting Endolysin

Clostridium difficile (*C. diff*) is a pathogenic gram-positive bacteria capable of causing nosocomial infections that in recent years have grown in severity and incidence (Monaghan et al., 2008). The outcome of sequencing a novel *C. diff* bacteriophage (ϕCD27) was the discovery, cloning and characterisation of a new endolysin (CD27L) (Mayer et al., 2008). This lysin was active against a wide range of *C. difficile* strains and subsequently has been structurally and functionally characterised (PDB ID 3QAY) (Mayer et al., 2011). Through structural manipulation its catalytic N-terminal domain alone has exhibited an increased lysis ability without greatly broadening the target bacterium range (Mayer et al., 2011). The structural and functional information available about this protein and the ability to manipulate its structure means it is an ideal target for functionalising Dps.

6.4.1.1 DpsA fused with full length CD27L

A gene coding for the CD27 lysin was synthesised *de novo* (Eurofins). The construct contained *Sac*I and *Not*I recognition sites either side of the CD27L gene allowing the gene to be removed from the pCR2.1 backbone via a *Sac*I/*Not*I digest and ligated into a *Sac*I/*Not*I digested pDpsA14 plasmid to create pDpsACD27L. This encodes the CD27 lysin translationally fused to the C-terminal end of DpsA. The histidine tag also remains in the vector at the C-terminal of the fusion for purification purposes. The total protein fusion has a molecular weight of 52.01 kDa, of which 30.15 kDa is attributed to CD27L. A dodecamer would thus have a mass of approximately 624.12 kDa. The predicted iso-electric point (pI) of the protein is 6.34 based on primary sequence.

Testing for soluble expression of DpsACD27L was carried out in transformed *E. coli* BL21 (DE3) using different induction times of 1 hour, 2 hours and an over-night induction; an uninduced sample was also included as a control. The temperature and IPTG concentrations were kept as previously described. Soluble expression was achieved under all conditions tested even the over-night induction sample which supplied the most amount of recombinant protein (Figure 6.7A). Purification of soluble DpsACD27L was achieved using IMAC; however not only were there issues with protein precipitation *in vitro* but also resolving the protein by BN-PAGE resulted in a smear (Figure 6.7B). The smear is indicative of non-specific protein aggregation. However with the smear stretching the length of the lane it also points to potential non- assembling proteins or protein degradation. Of course this result alone cannot be used to distinguish between problems in protein folding at secondary, tertiary and quaternary levels. There could be a range of reasons why problems were encountered in the conditions used and highlights that optimisation of the *in vitro* conditions will be required in order to achieve soluble recombinant protein.

6.4.1.2 DpsA fused with truncated CD27L (CD27Lt)

The problems encountered using the full length lysin can be caused by a range of issues for example non-specific interactions caused by the lysin or steric hindrance caused by the accumulation of three lysins located around the C-terminal pore region. Even though this would appear not to have been a problem with DpsAmCh, the CD27L is larger. In order to alleviate the problems seen, a fusion with a reduced size was also created and tested. The catalytic domain of CD27L at the N-terminal has previously been shown to retain functional activity, albeit with slightly less specificity (Mayer et al., 2011). Since this domain is smaller in size, it may result in a fusion that is capable of assembling without problems of aggregation and precipitation.

A truncated CD27L gene was created by amplifying the CD27L gene by PCR from the synthetic gene plasmid pCR®2.1CD27L using primers;

CD27LPCRF1 (5'-ATGAGCTCGATGAAAATTTGCATTACAGTGGGG-'3)

CD27LPCRR1 (5'-ATGCGGCCGCGACGCCTTCGTTATTAAT-'3).

The restriction sites for *Sac*I (CD27LPCRF1) and *Not*I (CD27LPCRR1) within the primers are underlined. The amplified fragment (566 bp) was digested with *Sac*I/*Not*I and cloned into pDpsA14 digested with *Sac*I and *Not*I. This plasmid was introduced into *E. coli* JM109 to bulk plasmid stocks and once correct cloning was confirmed by sequencing, a plasmid clone was transformed into *E. coli* BL21 (DE3). The truncated fusion reduces the monomeric molecular weight from 52.01 kDa to 42.24 kDa. This also reduces the potential dodecameric mass to 506.88 kDa, a significant 117.24 kDa less than the full-length fusion.

Soluble expression of DpsADC27t was explored on a small scale and whilst there was a large proportion of protein in the insoluble fraction, there was also a reasonable amount within the soluble fraction (Figure 6.7C). Expression and purification was carried out as previously described and resulted in the issue with protein precipitation being significantly reduced. Nevertheless, electrophoretic mobility via BN-PAGE was not improved (Figure 6.8A). Despite this, urea was tested to evaluate the ability to disrupt the assemblies, and also to refold the protein in a controlled manner, resolving any issues that may have occurred with protein folding. Whilst partial dissociation was detected using 8 M urea (Figure 6.8A) the molecular mass was still very large. In an attempt to disorder the assembly further, a series of temperatures were tested which can assist the urea unfolding process. Furthermore in order to disrupt disulphide bridges that may occur within the structure of the CD27L domain, a reducing agent (DTT) was also included. The different conditions displayed no change in ability to disrupt assembly with very little impact produced by the increased temperature. The reducing

conditions also had little effect suggesting that disulphide bridge formation does not contribute to the resistance of dissociation and denaturation by urea (Figure 6.8B). The removal of urea was carried out to establish if the slight dissociation and thus re-association would aid the formation of a complex that assembles, as it should. In addition, different ionic strengths were explored to find if this impacted on proper assembly. Figure 6.8C demonstrates that removal of urea did not result in DpsACD27Lt entering the gel regardless of the ionic strength in which the refolding process occurred. Taken together the results show that the functional group fused to DpsA can cause significant problems with protein folding or assembly yet the region (DpsA or functional group) responsible for the problems is unknown. Given the success with DpsAmCh, it highlights this approach does in fact work but may not be successful for all groups tested. Therefore, further functional groups can be explored to find if the platform is suitable for additional studies.

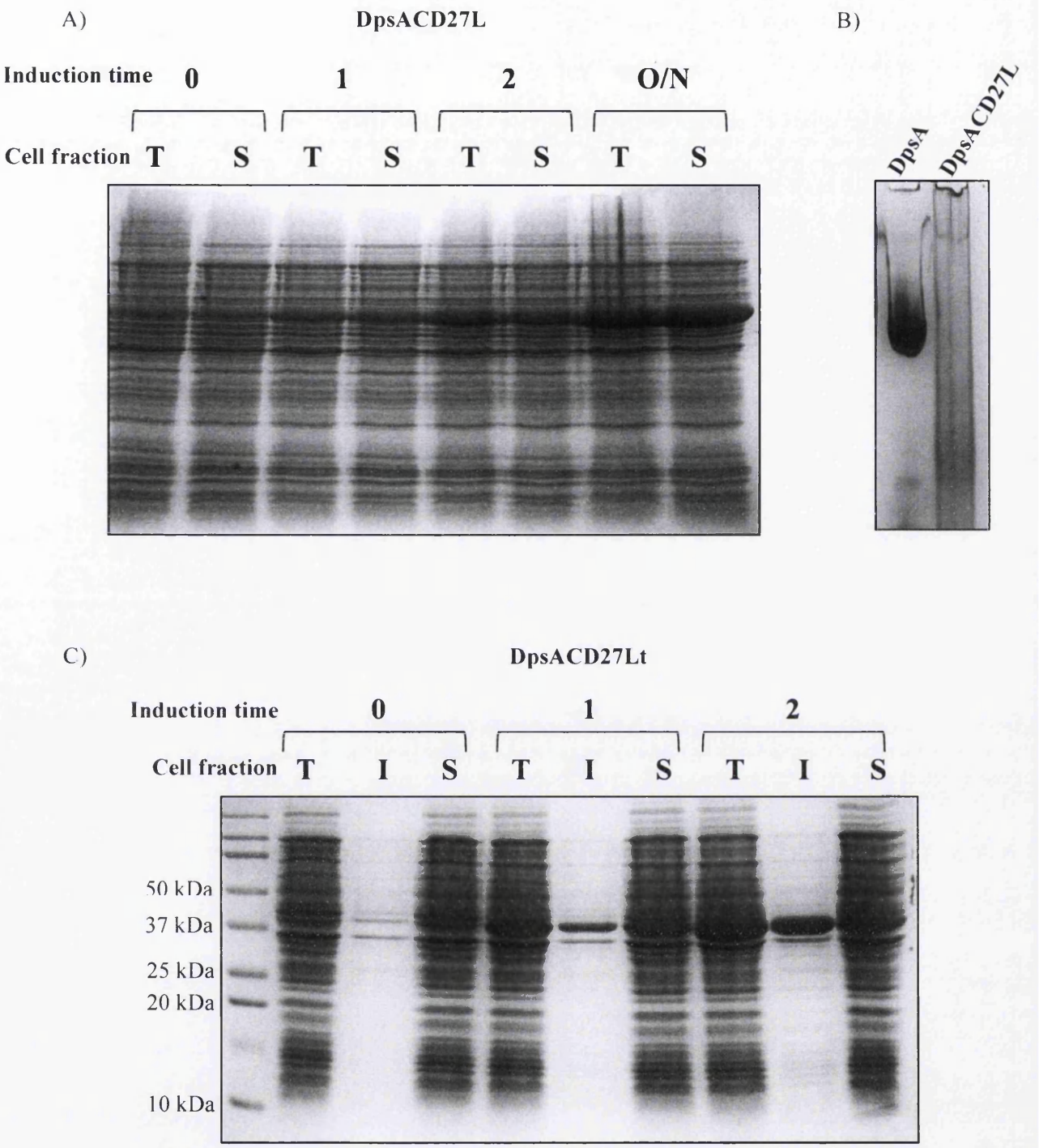


Figure 6.7: Expression analysis and oligomeric state analysis of DpsACD27L fusion (A) DpsACD27L. Gel analysis of crude total (T) and soluble (S) extracts from *E. coli* BL21 (DE3) before and after expression induction. Samples were taken at time point 0 hours, 1 hour and 2 hour. An overnight incubation of circa 15 hours was also included. (B) After purification pDpsACD27L was resolved by BN-PAGE but suffers from aggregation and precipitation. (C) DpsACD27Ltruncated soluble protein expression gel analysis of total (T), insoluble (I) and soluble (S) extracts taken at time points 0, 1 and 2 hours post induction. Biorad Precision PlusProtein™ Kaleidoscope standard was used.

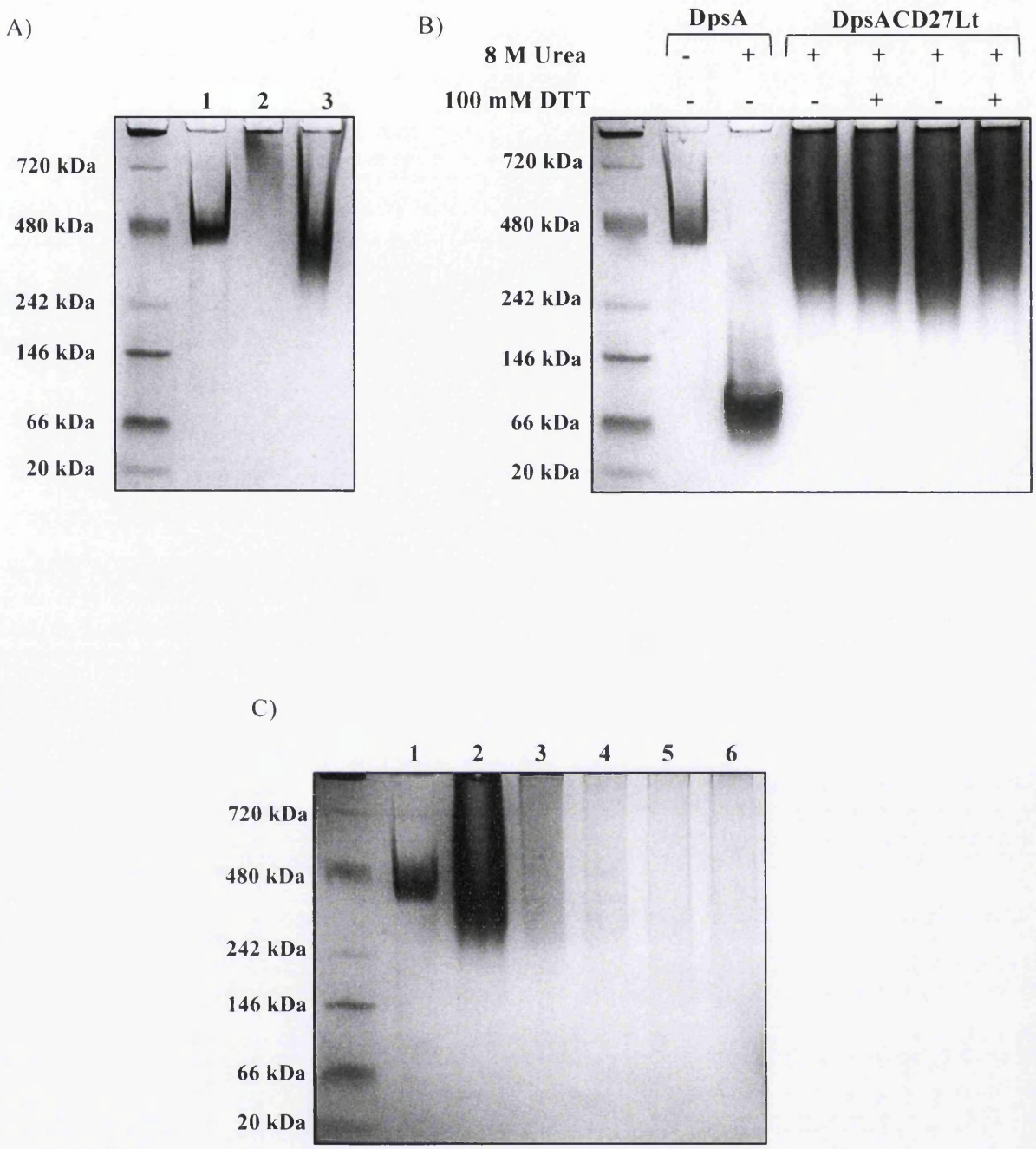


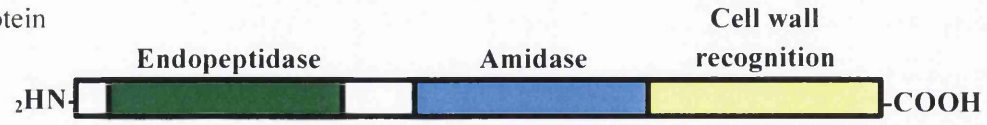
Figure 6.8: Assembly and dissociation analysis of DpsACD27Lt. Oligomeric assembly of DpsACD27Lt was analysed using BN-PAGE along with its ability to be denatured (A) lane 1 – DpsA, lane 2 – DpsACD27Lt, and lane 3 – 8 M urea DpsACD27Lt. (B) The optimisation of urea mediated dissociation failed to fully disrupt assembly of DpsACD27Lt even at 30°C and 37°C with and without 100 mM DTT. Conditions are described above each lane. (C) Removal of urea did not result in a DpsACD27Lt being capable of migrating into gel regardless of the ionic strength used. Lane 1 – DpsA, lane 2 – DpsACD27Lt 8 M urea, lane 3 – refolded DpsACD27Lt 200mM NaCl, lane 4 – refolded DpsACD27Lt 300mM NaCl, lane 5 – refolded DpsACD27Lt 400mM NaCl, lane 6 – refolded DpsACD27Lt 500mM NaCl. The protein standard used is the Invitrogen NativeMark™.

6.4.2 *Staphylococcus aureus* Targeting Endolysin

The *Staphylococcus aureus* bacteriophage φMR11 encodes a lysin (MR11L- also referred to as MV-L) that also displays specificity in targeting and lysing its host (Rashel et al., 2007). Multidrug-resistant *Staphylococcus aureus* (MDRSA) is another pathogenic bacterium posing an immense challenge to clinicians and scientists. It is considered of great importance that systems are developed to act as therapeutics and diagnostics as common antibiotics become increasingly ineffective. The MR11L can potentially fulfil both these roles; however, to be transformed into a diagnostic probe a means of tracing and detecting the MR11L is required. The Dps platform provides an ideal solution, which can be based on chimeric assembly of fluorescently labelled Dps with MR11L labelled Dps or simply by utilising the iron oxide paramagnetic core.

The MR11L contains three domains which consist of an endopeptidase at the N-terminal region, an amidase located within the central region of the protein and a cell wall recognition domain at the C-terminal region (Figure 6.9A) (Rashel et al., 2007). Studies previously conducted within the lab into functionalising DpsA with the full length MR11 lysin gave negative results showing an inability to obtain soluble protein from purifications using both native and denaturing conditions. Any further work using this lysin would thus require a different construct and so prompted the creation of a truncated MR11 lysin, which consisted of the cell-binding domain only. In order to create this truncated lysin, the full-length construct (pDpsAMR11L) was manipulated by site directed mutagenesis (SDM). SDM was used to introduce a *Sac*I recognition site down-stream of the amidase domain. Primers were designed for use with the QuikChange® Lightning Site Directed Mutagenesis kit (Figure 6.9B). Two bases were altered (underlined) which changed the sequence from 5'-GTTGCGAGTGCATGG-'3 to **5'**-GTTGCGAGCTCATGG-'3. Following SDM, the mix was transformed into *E. coli* JM109. Subsequent to the purification of possible mutant clones the plasmids were digested with *Sac*I; this results in the excision of a 386 bp fragment from the plasmid, which codes for the endopeptidase and amidase domains of the protein. The digested plasmids were purified from an agarose gel and religated in the absence of the fragment. The ligation mix was once again transformed into *E. coli* JM109 and colonies containing the putative mutant religated clones selected and grown for plasmid extraction. Sequence integrity of the plasmids was confirmed by sequencing using the universal T7 promoter and T7 terminator primers (LGC Genomics). A single correct clone was introduced into *E. coli* BL21 (DE3). The manipulation of the plasmid results in a genetic construct coding for DpsA fused at the C-terminal with the cell recognition domain of the MR11 lysin followed by the hexa-histidine tag (Figure 6.9C). The new DpsAMR11t monomeric mass is 32.99 kDa with a predicted pI of 5.5. The mass is immensely smaller than the full length DpsA-MR11L fusion that is 75.86 kDa.

A) Protein



B)

mutDpsALysF2
5'-GCTTCAAGTAATACAGTTAAACCAGTTGCGAG**ct**CATGGAAACGTAATAAATATGG-'3

mutDpsALysR2
3'-CCATATTTATTACGTTTCCATG**ag**CTCGCAACTGGTTTAACTGTATTACTTGAAGC-'5

C)

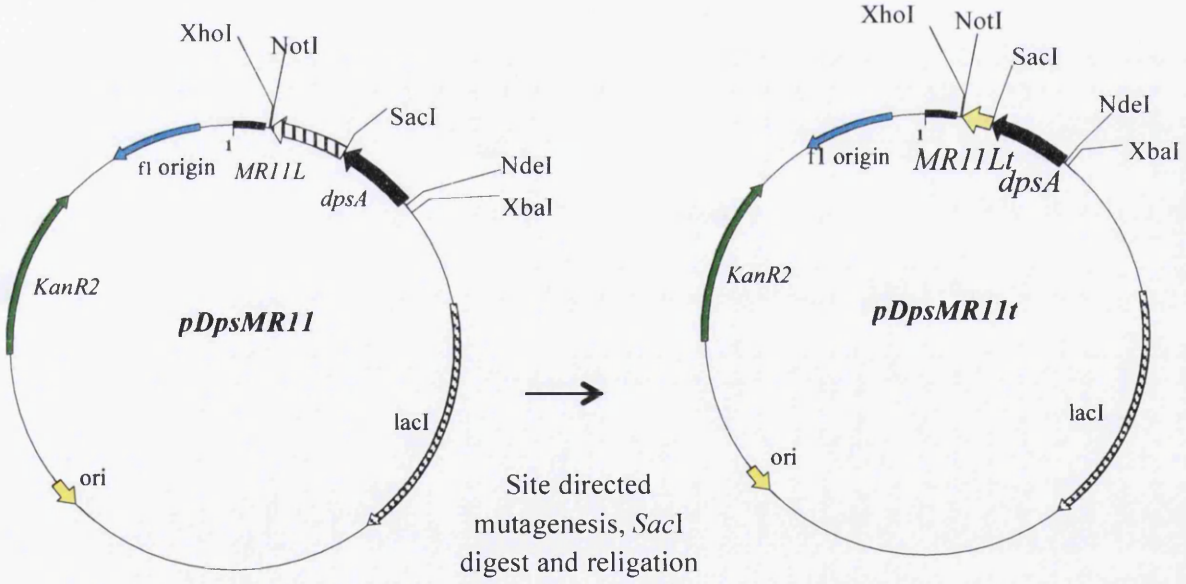


Figure 6.9: Construction of DpsAMR11t – A truncated endolysin fusion. (A) Schematic graphical representation of the domains of the MR11L protein sequence which include an endopeptidase domain, an amidase domain and a cell wall recognition domain. Below the line indicates the coding domain sequence (CDS) with the arrow representing the introduction of a *Sac*I site by SDM. (B) Forward and reverse primers used for QuikChange® SDM. Underlined in bold lowercase are the two mutations. (C) The cloning steps involved in creating pDpsAMR11Lt that includes the introduction of a *SacI* site by site directed mutagenesis, excision of the unwanted fragment and religation.

Following purification, DpsAMR11t assembly was assessed by BN-PAGE (Figure 6.10); the gel displays a substantial reduction in aggregates with clearer bands present albeit at masses larger than would be predicted for dodecameric assembly. The addition of 2 M urea improves electrophoretic mobility, which suggests that weak non-specific interactions may be the cause of problems with aggregation. This also suggests that protein stabilisation can be enhanced by the addition of co-solvents which can prevent aggregation. Of course this would require a great amount of buffer optimisation and is protein specific so would be required for each individual construct. Successful disruption of the assembly using 8 M urea highlights the potential that this construct has as a platform on which to develop the technology. There are further experiments that were not tested that such as hetero-oligomeric assembly and ferroxidase activity. However, the most important function to test would be the ability of the truncated MR11 lysin to selectively bind to *S. aureus* cells. These experiments would determine if this construct was suitable for development as a biotechnological application.

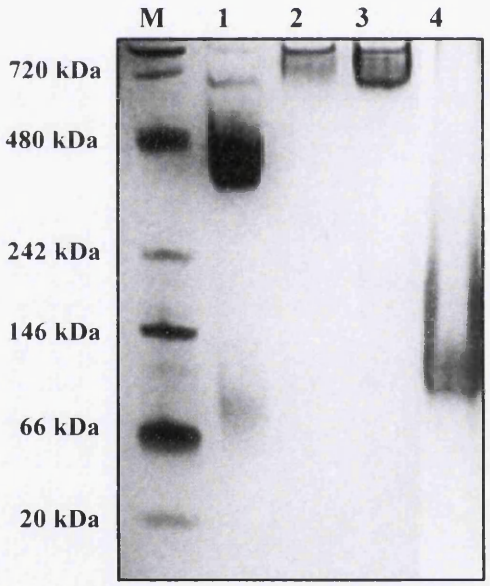


Figure 6.10: Assembly and dissociation analysis of DpsAMR11t. Composite image of BN-PAGE assessing molecular weight of DpsAMR11t truncated lysin with and without urea. Lane 1 – DpsA, lane 2 – DpsAMR11t, lane 3 – DpsAMR11t 2 M urea, lane 4 – DpsAMR11t 8 M urea. Invitrogen NativeMark™ was used as a standard (lane M). The addition of 2 M urea aids electrophoretic mobility, while 8 M urea dissociates assembly.

6.5 Summary of Exploring the Potential Nanotechnological Applications of DpsA Results

Nanotechnology is a term given to the study and manipulation of matter with a scale of 1 nm to 100 nm in at least one dimension. The *Sc*Dps proteins possess multiple structural features that are on a nanoscale; the external diameter of DpsA is around 10 nm and the internal cavity around 4 nm. As a naturally occurring ordered bio-assembly, DpsA is also capable of depositing iron within its internal cavity; it can be described as a size-constrained nano-reactor. These useful features prompted the exploration of DpsA with the purpose of studying its nanotechnological potential. Using the knowledge gained from the previous chapters, along with the available literature pertaining to the biotechnological exploits of ferritins and other Dps, it was conceived that DpsA would provide a good platform to investigate how these proteins can be manipulated into valuable biomedical applications. The key points that can be taken from this chapter follow:

- The oligomeric assembly of DpsA can be disrupted by 8 M urea and re-assembled following the removal of the urea.
- Formation of hetero-oligomeric assemblies between DpsA and non-dodecameric mutant variants can incorporate non assembling species into the protein cage.
- Hetero-oligomeric assemblies may facilitate a means of modulating and controlling assembly of dodecamers.
- Hetero-oligomeric assembly between DpsA and DpsB does not occur *in vitro* and thus is not likely to serve as an activity modulating mechanism *in vivo*.
- A fluorescently functionalised DpsA cage retains biological activity; ferroxidase activity with iron mineralisation and fluorescence of the mCherry fusion.
- Hetero-oligomeric assemblies can be formed between DpsA and a functionalised DpsA (DpsAmCh) creating a method to modulate the number of functionalised subunits per cage
- A truncated MR11 lysin showed improved prospects for functionalising DpsA with cell recognition proteins although functional activity was not tested.

These results do suggest that it is possible to functionalise a Dps nanocage and deposit an iron oxide core within the cage. Furthermore there were a number of constructs that were created in this study that were not actually investigated; full length and truncated lysins were also fused to DpsADTM. Since DpsADTM does not readily assemble into a dodecamer, this fusion may help to improve problems with solubility and aggregation. As the results have shown, hetero-oligomeric assembly can occur between DpsA and DpsADTM, therefore this process can be used to control when oligomeric assembly occurs, and if the process is

optimised, a technique of controlling the ratio of tagged to untagged subunits within the assembly can be obtained. Despite the potential, there are a great number of other issues that are likely to be functional group-dependent. Since mCherry worked as a fusion, and there was promise with the MR11 lysin, the problems encountered with the DpsACD27L and its truncated analogue suggests that the CD27L domain was the cause of the complications in achieving soluble, assembled and active protein.

The results as a whole underline the prospects for using Dps as platforms for the development of novel nanotechnology. Yet they also highlight the difficulties that can be encountered and the work that is required to overcome these problems.

Chapter 7

Discussion

7.1 Introduction

The Dps proteins, common to many bacteria and archaea, are key factors in the response to various stresses, both extracellular and intracellular in origin (Reon et al., 2012); (Boughammoura et al., 2012). Because of this Dps significantly influence colonisation abilities, stress resistance and persistence in many organisms (Colburn-Clifford et al., 2010); (Zheng et al., 2011). Dps generally share a quaternary protein architecture which involves the self-assembly of 12 subunits to form a spherical protein cage (Grant et al., 1998). The cage structure affords an environment suitable for the oxidation of ferrous iron with the internal cavity acting as a vessel for the formation and storage of a bio-available iron oxide core. This detoxification process nullifies Fenton chemistry and consequently serves as the main cellular protective feature of Dps (Zeth, 2012). Moreover, physical interactions with DNA and in some cases condensation and co-crystallisation also provide mechanical protection of the cell's genomic content (Wolf et al., 1999). Despite the structural assemblies that are largely conserved throughout the Dps family, not all proteins retain ferroxidase activity (Stillman et al., 2005), nor all have the ability to bind DNA (Ceci et al., 2003). However alternative and diverse roles are continually being revealed in a species- and environment-dependent manner; virulence in pathogenic bacteria (Tsuruta et al., 2012) or amide catalysis (Pesek et al., 2011) provide some examples. A growing interest in the nanoscale properties of the protein cages and their metal core fabrication capabilities means their continued study will not only provide further insight into the organism/structure/function relationship, but also may supply novel biochemical features that can be adapted for use as biotechnological tools.

To date, the study of Dps is largely confined to unicellular rod bacteria such as *E. coli*, however, the Dps of *Streptomyces coelicolor* represent an opportunity to explore the contribution of Dps to a complex filamentous bacteria. Furthermore, the unusual nature of encoding three *dps* genes poses its own questions with regard to the structural and functional differences that may be observed between the proteins. The work in this thesis has built on the findings that the *Sc*Dps proteins have expression patterns that are stress and protein specific. In addition, they have been implicated in nucleoid condensation during the reproductive growth phase of the host's life cycle (Facey et al., 2009). The "differential dual roles" of these proteins would suggest that their functional properties might be divergent; therefore the biochemical and structural analysis would reveal not only how these proteins differ, but also how they may fulfil their roles.

7.2 *Sc*Dps Proteins – Differential Assembly, Ferroxidase Activity and DNA Oxidative Protection

The *Sc*Dps proteins displayed variation in their ability to oligomerise to form dodecamers *in vivo* and *in vitro*. DpsA and DpsC were both capable of assembling into high oligomeric states, although the presence of a lower oligomerisation state of DpsA indicates that not all the protein may be associated with its dodecameric assembly. Indeed, this behaviour has been reported to occur in *E. coli* Bfr, although in this instance poor packing caused by water bridged asparagines at the four fold interface and a diminished dimer interface were demonstrated to be the causes of the "crippled assembly" (Ardejani et al., 2011). The inconsistent occurrence of the DpsA lower state between protein sample preparations throughout the work presented here would suggest that slight conformational variations may attain a similar effect to that encountered with *Ec*Bfr. On the other hand, analysis of the DpsA crystal structure revealed a dimer interface with structural characteristics highly similar to other characterised Dps. *In vitro* analysis of DpsA exposed it to a range of conditions that did not lead to disruption of the oligomeric assembly, thus the effects of ionic strength, slight pH fluctuations or temperature are unlikely to be the source of the incomplete assembly. However, protein degradation was not explored as a possible source. With the terminal tails implicated in oligomeric assembly of DpsA, loss of tail regions from a proportion of subunits may result in non-assembling subunits. In contrast to DpsA, DpsC preparations did not show lower oligomeric species throughout the work.

The functional properties of DpsA and DpsC presented additional differences between the proteins. Consistent with other Dps proteins DpsA has a preference for H_2O_2 as the oxidant for Fe (II) oxidation (Bellapadrona et al., 2009).The efficiency of iron oxidation was superior in the presence of H_2O_2 rather than O_2. The reaction was fast, reaching a maximum absorbance at the first reading following the addition of H_2O_2, a situation also described for the study of *B. anthracis* Dps (Dlp-1 and Dlp-2) (Liu et al., 2006). The rate of oxidation of Fe (II) by DpsA using oxygen could be monitored under our experimental conditions. DpsA delivers a more efficient oxidation process in the presence of O_2 in comparison to auto oxidation and provides a means of stabilising ferric iron in a soluble state. In the absence of the protein cage, the ferric iron precipitates out of solution. Reaction conditions were created where large quantities of Fe (II) were supplied to the cages in reactions that are not representative of conditions encountered under normal physiological circumstances. However this type of reaction represents the ability of the protein to continually oxidise iron in the presence of a growing iron oxide core. Iron was provided in excess; surplus to the number of FOC per cage. This ensures substrate limitation could not impact on oxidation rate. However this also manifests itself in reactions where the source of iron oxidation is not limited to the ferroxidase centres. Excess iron flowing into the

protein cage can become oxidised on the surface of the growing mineral core (Haikarainen and Papageorgiou, 2010), and this appears to occur in the reactions mediated by DpsA. Increased oxidation was seen in reactions where iron was already in excess per number of FOC thus it is likely that supplementary oxidation occurs simultaneous to oxidation mediated by the protein at the ferroxidase centres. The process of auto oxidation can be dismissed as a likely source of alternative oxidation since very little iron oxide precipitate could be visualised following the reactions, although this was not quantified.

It is likely that alternative oxidation was the factor that drove the increased oxidation observed with the DpsA deleted tail variants. Only DpsADCT supplied a cavity in which iron could be deposited. This species produces an oxidation curve indicative of a faster rate when compared to DpsA. Despite the complications with determining the oligomeric state of DpsADCT in solution, the fact that it binds ferric iron suggests that it maintains some form of cavity. This leads to the hypothesis that in solution the structure contains large openings through which iron gains rapid access to not only the FOC but also to the growing iron core, increasing the rate of oxidation. Furthermore DpsADTM produces a curve that continues to increase indicating that ferroxidase activity can be regenerated due to the absence of a cavity to fill, so saturation does not occur. The availability of unbound ferric iron within the solution can also favour and enhance the auto oxidation process contributing to the production of ferric iron consequently enhancing absorbance (Yang and Chasteen, 1999). Thus because DpsADTM does not sequester the ferric iron, it remains unbound in solution. DpsADNT has reduced ferroxidase activity, and while the reason for this remain unclear, accurate sizing and deduction of its oligomeric state may provide insight into how ferroxidation is impeded, possibly by obstruction of the FOC or electrostatic repulsion by an unusual multi-subunit arrangement.

Unexpectedly, DpsC precipitated upon addition of ferrous iron to the reaction mixture in conditions where DpsA managed to oxidise the iron. Oxidation was eventually achieved by sequential addition of reduced quantities of iron to the reactions subsequent to supplying the reaction with H_2O_2. The observed precipitation could be attributed to a "salting out" effect as seen when using ammonium sulphate precipitation. The competition for water between ammonium sulphate and hydrophilic amino acid side chains on the surface of the protein results in a decrease of protein hydration. This increases protein-protein interactions leading to protein precipitation. DpsC is also enriched with increased numbers of salt bridges particularly across the dimer interface; these features may provide additional stability to the protein in extreme conditions such as high temperature. However the screening effect of the double sulphate salts (ferrous sulphate and ammonium sulphate) may be enough to disrupt the bridges leading to dimer disruption and thus formation of insoluble protein salt. The ability to oxidise ferrous iron while simultaneously removing hydrogen peroxide allows DpsA and DpsC to protect DNA

against oxidative damage caused by Fenton chemistry. It also results in the formation of a ferrihydrite core within the cavity of the proteins. The positive results gained from the specific stain for protein bound ferric iron in polyacrylamide gels confirm the iron binding ability of DpsA and DpsC. It also strengthens the case that these assemblies observed in native PAGE are likely to be dodecameric.

7.2.1 DpsB presents itself as a Non-Dodecameric Dps

Under the conditions tested *in vivo* and *in vitro*, DpsB did not display large multimeric assemblies as demonstrated by an increased electrophoretic mobility of the protein when resolved via poly-acrylamide gel electrophoresis under non-denaturing conditions. This non assembling Dps is not a unique example; the *Dr*Dps1 protein's dodecameric assembly is only initiated under high salt conditions, with a dimeric assembly adopted in low ionic strength conditions (Grove and Wilkinson, 2005). Highly stable dimers that do not readily assemble to form dodecamers have also been discovered through mutational experiments where four individual residues regarded as hotspots in the *Ec*Bfr protein each removed the detectable formation of a 24-mer in solution (Zhang et al., 2010).

Corroborating the proposed dimeric assembly of DpsB is its capacity to exhibit ferroxidase activity, revealed by its capability to protect DNA against oxidative damage. In spite of the absence of quantitative analysis of DNA degradation, it would appear that DpsB has decreased protection efficiency when compared to DpsA and DpsC that originates from the lack of an internal cavity for the removal of the iron (III) from solution. Since ferroxidase activity is conditional on the formation of the dimer interface in order to supply the iron binding residues that generate the ferroxidase centres, ferroxidase activity of DpsB points to a dimeric state. Based on the electrophoretic migration of DpsB being parallel to that of BSA (66 kDa), an alternative oligomeric state would be predicted. According to this molecular weight, a trimeric association would be the likely state. Given that dimers are required for ferroxidase activity, the ability of DpsB to protect DNA against oxidative damage suggests that a dimer is almost certainly present. This rules out the formation of two of the trimer structures found within the dodecameric assembly ("ferritin-like" and "Dps-like"), however the formation of a trimer formed from a dimer and a single monomer is also unlikely due to the propensity for dimers to be a preferential and probable first assembled structure in the creation of a dodecamer. Accordingly it would suggest that the protein dimer might not migrate according to its true molecular weight.

Very little exploration into the assembly of DpsB was carried out. Despite the *in vivo* analysis suggesting that dodecameric assembly does not occur in its native environment, the influence of the sizing technique and chemistry of the solutions cannot be ignored. Although

DpsA and DpsC demonstrated multimeric assemblies and act as positive controls for the experimental set up, the oligomerisation of DpsB may be conditionally dependent and may require altered environmental chemistry; altered ionic strength, pH and presence of metals for example. This remains a facet of work that requires additional study.

7.2.2 Native PAGE Does Not Provide an Efficient Resolution for Determination of ScDps Oligomeric States

The use of native PAGE to predict molecular weight and permit calculation of the oligomeric state for these proteins proved to be ineffective. Both DpsA and DpsC displayed migrations that would be associated with a larger molecular weight than they have. Both proteins migrate closer to the apo ferritin band 1(480 kDa) rather than β-phycoerythrin (242 kDa). DpsA (268 kDa) was predicted to have a 387.03 kDa weight, corresponding to a 16-mer assembly, an oligomeric state very unlikely to occur. The respective crystal structures both indicated that DpsA and DpsC exist as dodecamers, and the tertiary structures both assume the Dps-like four helix bundle as opposed to a ferritin-like bundle, therefore removing the chance that the proteins assemble as 24-mers. Under the basic pH conditions of the gel and running buffer, the proteins, with acidic isoelectric points would carry a net negative charge; furthermore charge based migration was also aided by supplementation with a Coomassie G-250 additive which upon binding to positive regions confers a negative charge. The use of clear native PAGE for the iron staining procedure also revealed comparable migrations of DpsA and DpsC suggesting that the Coomassie additive is not necessary to confer a negative charge for migration. Separation through the gel matrix is not only based on the size of a protein, but also the shape and conformation of the proteins. Dps contain an empty internal cavity, which enlarges the dimensions of the proteins without increasing molecular weight. This provides another possible reason why their migration through the pores of the polyacrylamide matrix is retarded.

The Dps proteins explored in this thesis also involved non-dodecameric assemblies; DpsB and the DpsA deleted tail variants all had intermediate oligomeric states that were not dodecamers. As described in the last section, DpsB migrates consistent with it being a trimer, yet it retains ferroxidase activity, indicative of a dimer. Likewise, DpsADTM is predicted to be a trimer yet it also retains ferroxidase activity. The evidence of functional activity shown by these proteins would suggest one type of oligomeric assembly, while the gels suggest a larger oligomeric state. So it is likely that the dimers themselves also have migration profiles that are not true to their molecular weight, even in the absence of a cavity. Accordingly, the migration of high oligomeric state proteins is not only disturbed by an enlarged hydrodynamic

space/molecular weight ratio, but also the monomers themselves not accurately migrating according to their molecular weight.

It can be concluded that the non-gradient native PAGE approach used in this work does not have a good enough resolution to accurately assess the oligomeric assembly of these proteins, but works as an efficient tool to provide results that confirm whether dodecameric assembly occurs or not. Whilst native PAGE has been adopted as a tool in various reports to analyse the multimeric assembly of cage proteins (Ardejani et al., 2011) (Chowdhury et al., 2008) (Roy et al., 2007), it is almost always accompanied by higher resolution techniques such as gel filtration chromatography, analytical centrifugation or electron microscopy. These separation and sizing techniques remove any influence the chemistry of the PAGE system has on the oligomeric state and often use experimentally determined equations for estimation of molecular weight that do not rely on calibration methods and assumptions about protein shape.

7.3 Role of Tails in Oligomeric Assembly

The tails of various Dps proteins have been described to participate in maintaining and controlling oligomeric assembly. This study has explored the roles of the tails of DpsA and DpsC by protein manipulation studies by assessing oligomeric assembly upon removal of the tail structures and also through analysis of the crystal structures. Although the analyses were conducted independently, the results ultimately converged to indicate the importance of the tails to DpsA and DpsC assembly. In addition, the DpsADTM protein, lacking both N and C tails displays assembly features similar to DpsB and suggests that some Dps require a tail for some form of assembly.

7.3.1 The N-terminal Tails of DpsA and DpsC – Importance in Oligomeric Assembly

The N-terminal tails of DpsA and DpsC were both found to be vitally important to the oligomeric assembly of the proteins. An aggregated oligomerisation process was caused by the fusion of a histidine tag to the N-terminal region of DpsA and provided the first evidence that suggest a role for the N-tail in assembly. Furthermore, the deletion of the N-terminal tail resulted in loss of efficient oligomeric assembly when in solution. The DpsADNT mutant yields a single low oligomeric state that fails to form a dodecamer and has a more drastic effect than the loss of the C-tail, which, in the case of DpsADCT, preserves some form of high order complex. The assembly of DpsC also becomes massively distorted without its N-terminal tail. Almost mirroring the N-tagged DpsA, DpsCDNT forms soluble aggregates that do not enter the

native PAGE gel, but instead remain within the wells. This makes it difficult to assess if any form of self-assembly actually occurs, a dimer for example, or if it is all purely aggregation.

The crystal structures of DpsA and DpsC revealed how each N-tail contributes to the stabilisation of dodecameric assembly, but also provide specific interactions that can explain the differences observed in dodecamer stability. The DpsA N-tail was greatly involved in the "ferritin-like" trimer interface, also providing a slight contribution to the dimer interface. Each N-tail at the "ferritin-like" interface of DpsA overlaps the C-terminal region of a symmetry related subunit four helix bundle; interacting with its C and D helices in addition to the C-terminal tail. The linking of N and C termini highlights the importance of each tail to assembly. However the influence that this interaction has on the conformation of the tails cannot be inferred from this work. A single residue left of the C-tail (Glu-162) in DpsADCT mutant was found to be responsible for maintaining the interaction with the N-tail. The preservation of an N to C-tail interaction in the deleted C-tail mutant meant that the N-tail adopts the same conformation as in full length DpsA structure, thus it is impossible to deduce if this conformation is due to the interactions with the C-tail or interactions with the adjacent subunit C and D helices. The importance of DpsA's N-terminal tail is accentuated by Virtual Alanine Scanning results which predict Lys-8 and Tyr-9 to be potential hotspot residues; identifying them as being energetically important to the preservation of the "ferritin-like" interface. Despite five other residues contributing to the "ferritin-like" interface, it would be interesting to see the effect of *in vitro* mutation of these two hotspots to see if this would be enough to emulate the effects observed in DpsADNT.

DpsC has an N-tail that is substantially longer than that of DpsA. Aside from the *D. radiodurans* and *K. radiotolerans* Dps which contain helical segments within their N-tails, the DpsC N-tail equals the length of resolved N-tail residues from Dps structures in the PDB, matching the *T. elongatus* Dps at 17 residues resolved within the structure. However the DpsC N-tail has a further 27 residues disordered. The crystal structure confirmed that tail residues were involved with interactions across the surface of the dodecamer, acting as a brace to the dimer interface at each end of the anti-parallel dimer. The tail also assumes a role at the "ferritin-like" interface similarly interacting with C-tail residues. Its importance at both of these interfaces provides an explanation for the impaired high order assembly of DpsCDNT. While investigation into the ferroxidase activity of DpsCDNT was not carried out, it would provide evidence to establish if the loss of the N-tail causes aberrant dimer formation. The importance of the tail to both the dimer and "ferritin-like" interface is also defined by the N-tail having a residue in each interface that is considered a hotspot for the stability of the interface. Ile-30 is considered a hotspot within the dimer interface of DpsC; in addition this residue also contributes to the "ferritin-like" interface although without being regarded as a hotspot. Only

one N-tail residue from the "ferritin-like" interface is considered a hot spot, although this is not surprising given that DpsC would appear to place emphasis on stabilising the dimer interface and "Dps-like" interface.

7.3.2 Contribution of the C-tail to Assembly of DpsA and Prediction of DpsC C-Tail Influence on assembly

Experimental evidence was able to prove that the C-tail of DpsA was also important for the self-assembly and stability of DpsA, although based on our predictions, a different structure is formed when in solution and within the crystals. In solution DpsADCT displays a migration through native PAGE that would correspond to a non-dodecameric assembly. It maintains ferroxidase activity showing it forms at least the dimer interface required furthermore it is capable of depositing iron, showing a cavity is formed. The increased Fe (II) oxidation when compared to DpsA may suggest that an open cage structure is being formed such as the decamer observed in a deleted tailed *M. smegmatis* Dps (Roy et al., 2007). This may facilitate an increased movement of iron to the ferroxidase centres; it may also allow oxidation to occur on the growing iron oxide core. The crystal structure revealed a dodecameric assembly extremely similar to DpsA (RMSD 0.387 Å between 158 atoms) and so this was unable to confirm the architecture of DpsADCT in solution.

The crystal structure of DpsA can still be used to provide evidence as to why this unusual assembly was seen and why the loss of the C-tail is not as detrimental as loss of the N-tail. The "Dps-like" pore has a higher complexation significance score than the "ferritin-like" pore of DpsA, thus it would be intuitive to suggest that its disturbance would lead to a greater disruption of oligomeric assembly. However, this interface is reliant on hydrophobic interactions that occur mainly between the four-helix bundles. Thus the loss of the C-tail may only weaken this interaction. The lack of predicted hotpots in the C-tail also supplements the idea that the N-tail is more important for maintaining assembly. On the other hand, the role of the Glu-162 residue in the assembly process cannot be disregarded. Even though it constitutes the first residue of the C-tail, it was included in the DpsADCT mutant so that the D helix was not disturbed. However, this residue is involved in both the "ferritin-like" and "Dps-like" interfaces; even though it is not considered a hotspot it is involved in interactions with the N-tail and so it would be interesting to see the effect of its removal on assembly.

Whilst an *in vitro* exploration into the role that the C-tail of DpsC plays in assembly was not undertaken, the structure reveals that it is likely to be vital to assembly. Only 8 residues long, the C-tail is likely to be important at the "Dps-like" interface. Triangulating the symmetry related subunits; the three tails brace the adjacent subunits and act to reinforce the interface. The PISA analysis suggests that DpsC possesses an unusual "Dps-like" interface, which is

considered the most significant for assembly in DpsC, displaying characteristics shown by Dps from extreme environments. The exact contribution of the C-tail to the interface stability remains hard to predict without *in vitro* evidence, particularly since the interface places a lot of emphasis on hydrophobic interactions between the helical bundles. On the other hand, alanine scanning indicates three hotpot residues (Leu-197, Val-198 and His-199). These residues represent ideal candidates for *in vitro* mutagenesis to explore the bearing that the tail has on dodecamer assembly and stability.

7.4 PISA Analysis Suggests an Extremophilic Like Stability of DpsC

An interesting facet of this work is the evidence that points to DpsC possessing "extremophile-like" properties. *In vitro* evidence points toward DpsC being endowed with enhanced dodecameric stability when compared to DpsA. Additionally, structural analysis of DpsC revealed that stability-enhancing features present within the protein's architecture could be the foundation of the observed stability.

Data from PISA analysis on a range of Dps proteins revealed a major divide in structural architecture permitting grouping into two domains; Dps from mesophilic organisms such as *Sc*DpsA, *E. coli*, *M. smegmatis* and *L. innocua* and Dps from extremophiles such as *T. elongatus, S. solfataricus* and *H. salinarum*. These extreme Dps are often characterised by an extended dimer interface area. This interface is also enriched with high frequencies of inter subunit hydrogen bonds and salt bridges which are thought to contribute to the protein stability differential between thermophilic and mesophilic species (Kumar and Nussinov, 2002). The extended dimer interface, abundance of salt bridges (19 per dimer interface) and hydrogen bonding interactions (23 per dimer interface) is thought to be the reason for the high thermostability of the *Ss*Dps (Wiedenheft et al., 2005). Furthermore another thermophilic Dps, *Te*Dps, also displays an increased "Dps-like" interface CSS score (1.000) when compared to the low values of mesophilic Dps (*Ec*Dps – 0.238) (Franceschini et al., 2006). It was quite remarkable to find these features in DpsC, where an extended dimer interface, containing 27 hydrogen bonds and 27 salt bridges, was also combined with an increased "Dps-like" interface area that holds the most negative solvation free energy gain upon formation of all the "Dps-like" interfaces from the proteins analysed. This contrasts the analysis reported for the mesophilic *E. coli* Dps which does not contain significant hotspots located at this interface due to the lack of salt bridging interactions or aromatic residues present (Zhang et al., 2011). On the contrary, DpsC possesses four aromatic residue hotspots at the "Dps-like" interface (Phe-81, Tyr-82, Trp-186 and Phe-187), which strongly suggests that DpsC has evolved a stabilising role for the divergent "Dps-like" interface.

In parallel to the work described within this thesis, phylogenetic analysis has revealed the protein sequence of DpsC to be very rare when compared with the available Dps-like sequences in databases (Facey et al., 2013). A phylogenetic representation of all Dps sequences available resulted in three distinct clades populated with "DpsA-like", "DpsB-like" and "DpsC-like" sequences. A schematic representation of this is provided in Figure 7.1. The "DpsC-like" clade contains a limited number of sequences, moreover it is populated by some actinobacterial species which have been demonstrated to originate from extreme environments; *Thermobispora bispora* for example (Liolios et al., 2010). Within these sequences, there are a lot of conserved residues both unique to these sequences and also conserved among Dps as a whole. Perhaps the most interesting conserved residue among the DpsC-like species is Arg-118. Located on the BC helix, Arg-118 forms a number of hydrogen boning interactions with Val-126 and Arg-128 and yields the highest $\Delta\Delta G$ complex for all the DpsC hotspots. This points to the importance of this residue for dimer stability, and is an ideal target for mutation of the BC helix. We can conclude that despite an unclear *in vivo* role, DpsC is a protein that displays extremophilic-like structural, biochemical and ancestral properties that suggests its acquisition into the genome of mesophilic *Streptomyces coelicolor* is likely to have occurred via horizontal gene transfer.

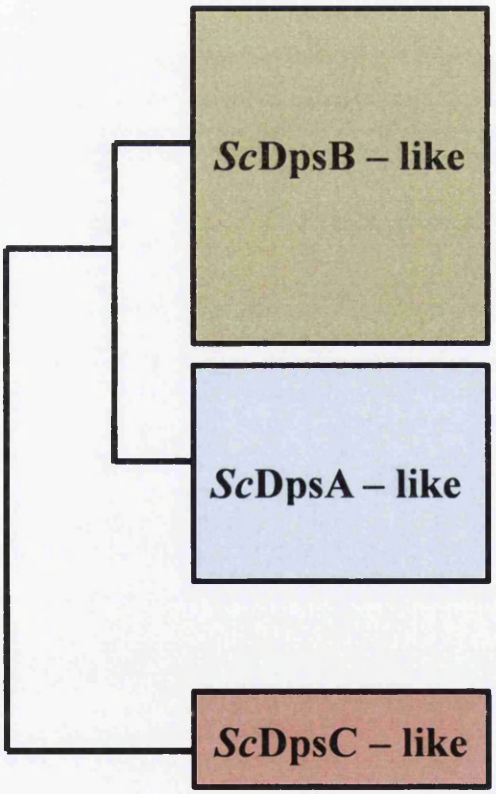


Figure 7.1: Schematic representation of phylogenetic relationship of *Sc*Dps proteins (adapted from (Facey et al., 2013). The limited number of DpsC-like proteins and their thermophilic origin suggest a unique phylogenetic history for DpsC; mesophilic organisms acquiring a "thermophilic adapted" gene.

7.5 DNA Interaction Was Not Detected but Require Further Investigations

The interaction between Dps and DNA has been postulated to occur in multiple ways; a direct interaction with the positive residues within the tail regions is observed in multiple Dps (Grant et al., 1998); (Ilari et al., 2000); (Roy et al., 2007), or indirect interaction is dependent on divalent metal ions such as Mg^{2+} bridging negative charges between the surface of the protein and the DNA (Frenkiel-Krispin et al., 2001). The experiments explored within this work concentrated on direct interactions that appear to be the most frequent type of interaction between Dps and DNA. No band shifts were detected in the electrophoretic mobility shift assays (EMSA) within this work, suggesting that under these conditions the *Sc*Dps proteins do not possess the ability to interact with DNA. Furthermore, the oxidative protection assays did not reveal any band shifts suggesting that Dps-DNA interaction was not triggered by the presence of Fe^{2+} and H_2O_2 as required for the *Campylobacter jejuni* Dps-DNA interaction (Huergo et al., 2013). There are a number of reasons why this does not provide conclusive evidence to discard DNA binding as a functional trait of the *Sc*Dps proteins, but if interaction does occur, it may not be achieved by a direct Dps-DNA interface. One possible explanation is a requirement for specific DNA topologies for the interaction to occur, and such topological conformations cannot be recreated in the *in vitro* conditions used. The cooperation of other *S. coelicolor* DNA binding proteins, inducing conformational changes in the DNA and promoting a putative DNA-Dps binding event cannot be overlooked.

If consideration is given to the tail regions, it is possible to identify regions that may or may not be involved in DNA interaction. Associated with a direct interaction is a flexible tail composed of multiple basic amino acids. DpsA contains a single positive residue within an N-tail that in all monomers is likely to be fixed in a single adopted conformation due to interactions with regions of adjacent subunits such as the C-tail. The importance of the DpsA N-tail to oligomeric assembly also emphasises its role in stabilising the structure, thus making it unlikely to mediate a DNA interaction. This characteristic is consistent with the reports within the literature where proteins such as the *Agrobacterium tumefaciens* Dps, whose N–tail is immobilised on the surface of the dodecamer, fails to bind DNA (Ceci et al., 2003). The C-tail of DpsA, also fixed in position by interactions on the surface of the dodecamer, contains no positively charged residues thus this is also not likely to associate with DNA. DpsC actually has a number of positively charged residues within the N-terminal tail. There are 4 Lys residues and 2 Arg residues. This N-tail is characterised as having the first 27 residues disordered within the structure suggesting a highly flexible region that may protrude into the solution. Within this region, 3 lysine residues and a single arginine residue provide the positive potential that is required for direct DNA interaction thus can be identified as a putative DNA binding region. Similar tail flexibility has been reported to be responsible for DNA interaction by the *D.*

radiodurans Dps1 where 30 structurally disordered N-tail residues mediate the interaction. Furthermore, due to the fact that this region is rich in lysine residues (6 in total), it is hypothesised that these residues confer the binding ability (Bhattacharyya and Grove, 2007). Conversely, the 4 positive and flexible residues of the DpsC N-tail are accompanied by 3 acidic residues, which may act to shield the positive charges generated, potentially repelling the negatively charged DNA. The C-tail of DpsC is adhered to the surface, and does not contain many charged residues thus is highly unlikely to be a source of DNA interaction.

There are a number of conditions that have been shown to impact on the ability of Dps to interact with DNA. The inclusion of divalent ions (Frenkiel-Krispin et al., 2001), pH of solution (Ceci et al., 2005) and ionic strength of *in vitro* reaction mixtures (Huergo et al., 2013) have all been demonstrated to modify the ability of Dps to bind DNA. Furthermore, the size and conformation of DNA has been implicated in the ability, affinity and molar ratio of DNA binding. The two lactococcal Dps proteins both display an absence of DNA binding ability in the presence of fragments 300 – 3000 bp in length. Nevertheless, when titrations were carried out using larger fragments (~8000 bp), DNA binding was observed (Stillman et al., 2005). In addition, the *Ec*Dps has been shown to display a differential Dps: DNA binding ratio which is suggested to be dependent on DNA conformation; linear vs supercoiled (Jeong et al., 2008); (Ceci et al., 2004). From this it is possible to suggest that exploration of all these conditions is required to confirm if the *Sc*Dps proteins retain a DNA binding ability *in vitro*. As it remains, the *in vivo* effect on *S. coelicolor* nucleoids caused by disruption of *dps* genes cannot be explained by direct interactions with DNA. However the presence of other intracellular proteins and cofactors that contributes to the architecture of the nucleoids *in vivo* or confers other forms of DNA protection may affect the ability of the *Sc*Dps to interact with DNA.

7.6 Hetero-oligomeric Assembly

Ferritins display an ability to form chimeric assemblies between H and L chains in order to modify their functional traits. In light of this, it was considered that the three Dps from *S. coelicolor* may adopt this approach as a means of modulating their activities *in vivo*. The fact that dodecameric assembly of DpsB was not detectable in recombinant or *in vivo* preparations of protein hinted that its incorporation into DpsA and DpsC assemblies might serve as its *in vivo* role. The approach used to test hetero-assembly is based on the reversible disruption of dodecameric assembly to a dimeric state; this is followed by renaturation/reassembly upon removal of the denaturant. Mixing of alternative protein subunits at the assembly-disrupted phase allows for the incorporation of mixed species into the dodecamer. This procedure was shown to successfully reverse the disruption of DpsA assembly alone; however addition of

DpsB to DpsA did not result in any detectable inclusion of DpsB into the dodecamer. Since the dimers are the building blocks of the dodecamers in this experiment, the only interfaces that form between DpsA and DpsC are the trimer "ferritin-like" and "Dps-like" complexes.

Analysis of the crystal structure revealed that in order for these interfaces to form, suitably positioned residues are required from each subunit in order for the interface to occur. Thus DpsB may not have the relevant residues to produce the "ferritin-like" or "Dps-like" interfaces. This is also expected to be the case for DpsC. However the fact that DpsC cannot be disrupted by urea meant that this could not be investigated. The divergence in sequence shown by these proteins and the difference in the interface characteristics between DpsA and DpsC, evidenced by PISA analysis, is enough to predict that hetero-oligomeric assembly between the *Sc*Dps proteins is not likely to occur. Hetero-oligomeric assembly was shown to occur between DpsA and its tail-less variants. The inclusion of DpsADCT and DpsADTM into dodecamers can be used to substantiate the claim that the tails involved in the "ferritin-like" and "Dps-like" interface are vital to stabilise the complex. In the case of DpsA, both N and C tails are required to brace the trimer interfaces as demonstrated by the fact that DpsADCT on its own cannot form a dodecamer, and requires the contribution of the N-tail from DpsA to form a dodecamer. Unfortunately, experimentation using DpsADNT was not carried out. Whilst it is likely that it can be incorporated into a DpsA/DpsADNT chimeric assembly, it would be very interesting to find if a DpsADCT/DpsADNT can be used to form a dodecamer, and this would provide additional evidence to prove that both tails are required for assembly.

These observations lead to a question: How many full-length subunits are required to incorporate tail-less species into a dodecamer? As the minimum building block of the assembly is a dimer, molecular modelling can be used to speculate that a minimum of two full-length dimers could be sufficient to stabilise a dodecameric assembly. These dimers provide a total of 8 terminal tails; 4 N-tails and 4 C-tails. Each N-tail can interact with a single DpsADTM monomer, while the C-tail can interact with a separate DpsADTM monomer, providing stability for a total of 4 DpsADTM dimers. Figure 7.2 can be used to depict how the assembly is stabilised by the tails; however it does not infer the actual stages of the self-assembly process. Placing the DpsA dimers (red) at the top and bottom, a DpsADTM dimer can be placed in between (white). Interfaces that form between these DpsADTM subunits and the N-tail of the DpsA subunits (green sections) are strengthened by intersubunit hydrogen bonding interactions; Thr-10 of the N-tail to Arg-158 of the DpsADTM four helix bundle, and Pro-12 of the N-tail to Arg-126 of the DpsADTM bundle (Figure 7.2A). Because of the anti-parallel nature of the dimers and the symmetry that exits within the dodecamer, these interfaces also occur on the opposite side of the assembly between the same DpsA dimers and a different DpsADTM dimer (Figure 7.2B). The addition of a further two DpsADTM dimers (pink) fill the remaining spaces

in the assembly and are held in place by hydrophilic interactions between Glu-162 of the C-tail from DpsA and Arg-104 located on the BC to C loop of the DpsADTM subunit. These represent the interactions that occur at the "Dps-like" interface (cyan sections) (Figure 7.2C).

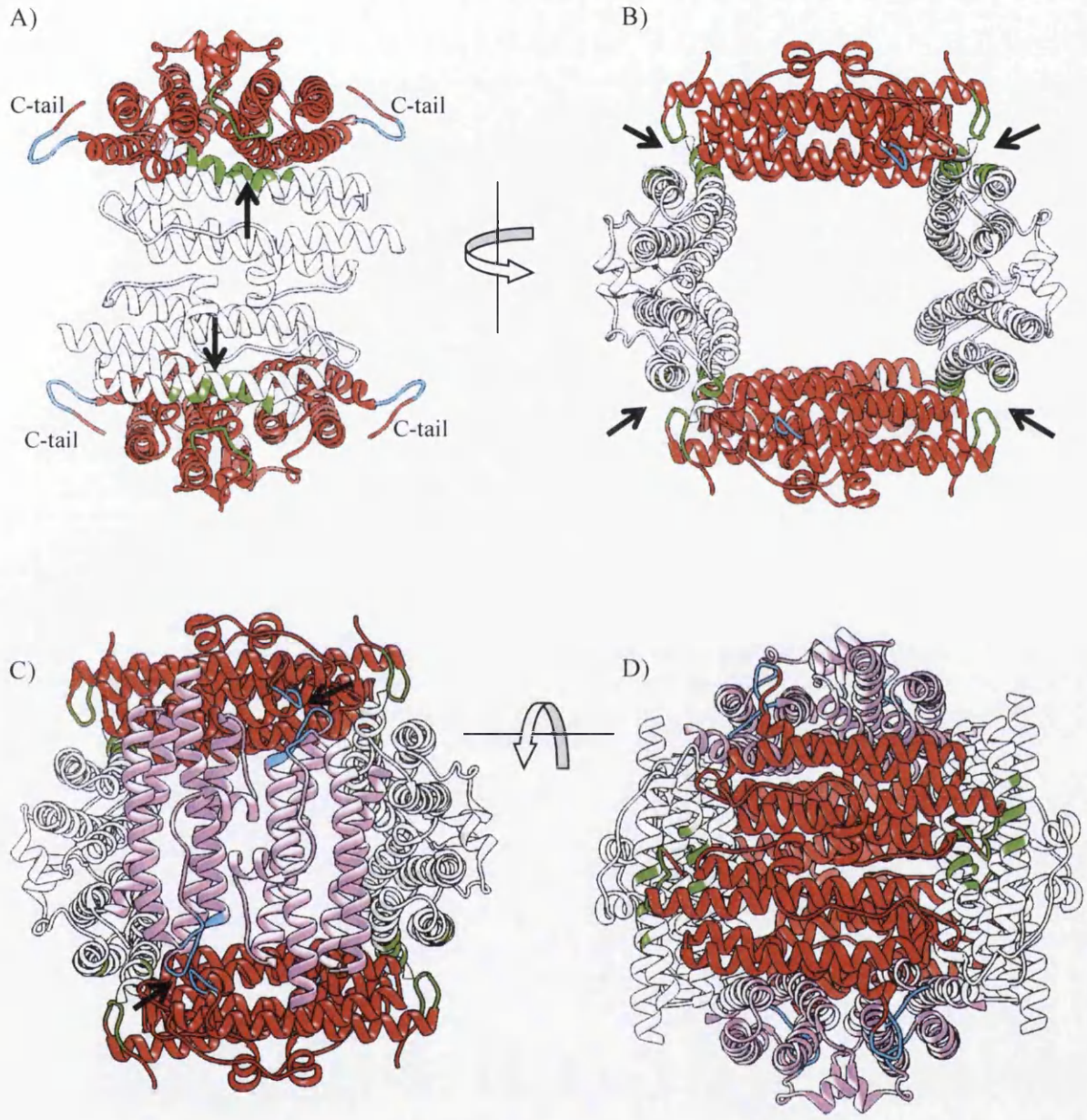


Figure 7.2: The minimum number of full length DpsA dimers required to incorporate DpsADTM into a dodecamer can be represented using a molecular model of the assembly. Two DpsA dimers (red) are required to provide tails, which stabilise a total of 4 DpsADTM dimers (white and pink). (A) Highlighted by the arrows are the N-tails of DpsA interfacing with the helix bundles of a DpsADTM dimer (green region). (B) The symmetry of the assembly shows how the interfaces with the N-tails occur at both sides of the structure. (C) A further DpsADTM dimer (pink) interfaces with C-tails of DpsA (cyan region). (D) The dodecamer is completed with a fourth DpsADTM dimer (pink) which also interfaces with the C-tails of DpsA. All the interfaces are strengthened by inter-subunit hydrophilic interactions between the DpsA tails and the DpsADTM helical bundles. Images were created using UCFS Chimera 1.7rc.

A set number of dimer ratios can occur within a chimeric assembly between DpsADTM and DpsA as long as a minimum of 2 full length dimers are present (Table 7.1).

Table7.1 Ratio of DpsA to DpsADTM in dodecameric assemblies

Number of DpsA dimers	Number of DpsADTM dimer
2	4
3	3
4	2
5	1
6	0

Remarkably when observing the multiple bands observed after resolving the DpsA/DpsADTM hetero-oligomeric assembly via native PAGE, five bands are evident, indicating that these five assemblies all occur (Figure 7.3). No work was carried out to drive the formation of a particular ratio, however adding an excess of either protein at the mixing stage may drive the formation of dodecamers enriched for particular dimer subunits.

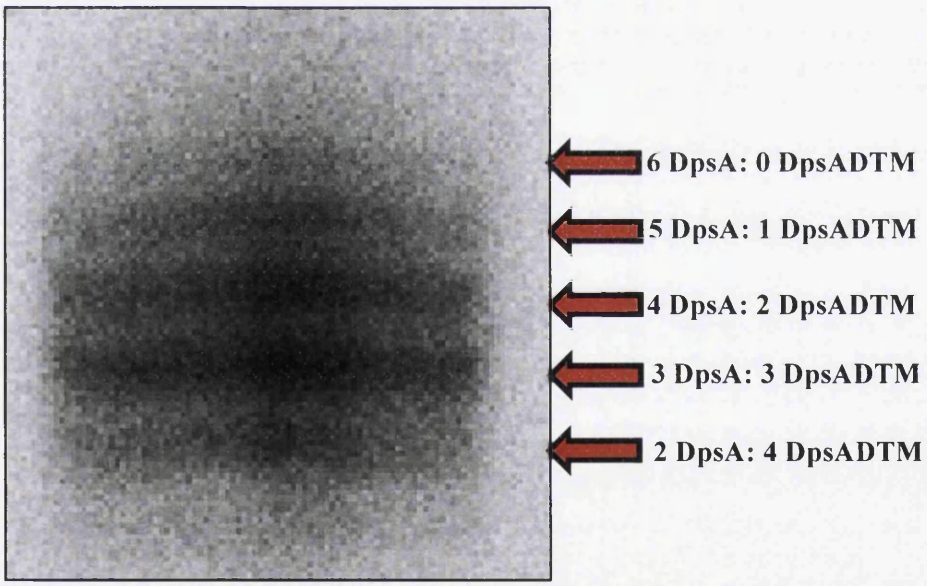


Figure 7.3: Native PAGE resolved DpsA/DpsADTM hetero-oligomeric assembly displays 5 different bands indicative of the ratios of DpsA to DpsADTM dimers within the dodecamers. Displayed besides each arrow is the predicted ratio. DpsADTM does not form a dodecamer therefore there is no 6:0 dimer ratio in favour of DpsADTM.

The ability to take a non-assembling Dps (DpsADTM), and incorporate it into a dodecamer provides a technique of manipulating assembly that has importance when using these proteins as a platform to develop nanotechnological tools. Not only does it supply a means alleviating problems caused by assembling 12 functionalised subunits within the self-assembling dodecamer, but also a method of multi functionalising the protein cage with different functional groups on the surface of the dodecamer.

7.7 ScDps as Nanotechnological Tools

Ferritins have been exploited as platforms for the development of several biotechnological applications with the focus based on manipulating the two distinct protein surfaces; the exterior cage surface and the interior cavity surface. These structures have been adapted by genetic manipulation or by chemical modification in order to impart a new or improved functionality to the protein such as adaptation of the internal surface for the preparation or encapsulation of novel inorganic metallic nanoparticles (Kramer et al., 2004). The external surface is often reserved as a template for the presentation of targeting ligands in order to fabricate a cage with specific recognition that can be detected or directed by its internal metal core (Gupta and Gupta, 2005). The type of functional group will depend on the end application but range from gold binding peptides for bio-template fabrication on gold electrodes (Ishikawa et al., 2008), to the integrin recognising peptide RGD-4C for the melanoma cell specific targeting of ferritin (Uchida et al., 2006).

The DpsA protein provides relevant traits that can be adapted for various nanotechnological applications. Producing a self-assembling protein cage with a diameter of ~9 nm, it contains a hollow internal cavity which is the likely receptacle of an iron oxide core generated by the protein's inherent ferroxidase activity. Because of these characteristics and the ability to disrupt oligomeric assembly, the DpsA cage was functionalised using a translational fusion approach through genetic modification of the coding sequence. DpsA was initially functionalised with a fluorescent protein (mCherry). The resulting hybrid fusion retained the ability to self-assemble into a bioactive cage capable of depositing an iron (III) core despite the functionalization of each of the 12 subunits. Whilst the dimensions of the assembly remain indefinite due to the resolution limitations of Native PAGE, its ability to oxidise ferrous iron demonstrates that the dimer interfaces are likely to be unaffected by the addition of the fusion. Moreover, the highly fluorescent signal detected from the proteins resolved by native PAGE also verifies that the fusion remains biologically active. However, this is not the first example of a fluorescent protein cage. Fluorescent ferritin nanoparticles (FFNPs) were created with both green and red tags and once conjugated with a DNA aptamer were used as a probe for the

detection of the cancer marker PDGF-BB (Kim et al., 2011). This construct showed an enhanced signal when compared to the eGFP alone and suggests that high numbers of fluorescent subunits in a confined area strengthens the fluorescent signal.

In order to create a functionalised particle with biomedical applications, DpsA fusions were created in which endolysins were supplied as the functional groups. Unfortunately these constructs did not prove to be as successful as the DpsAmCh fusion. The *Clostridium difficile* targeting endolysin CD27L lysin offers the ability to specifically and selectively target and lyse *C. difficile* cells. However both full length and truncated lysin constructs fused to the C-terminal of DpsA resulted in the propensity for the protein to aggregate and precipitate *in vitro*. There are number of reasons why these constructs did not behave as well as the DpsAmCh construct. At 52.01 kDa, the DpsACD27L monomer is ~ 4 kDa larger than DpsAmCh thus creating a greater chance of steric hindrance blocking the self-assembly of the DpsA subunits. However, truncation of the CD27L lysin to preserve only its catalytic domain reduced the monomeric molecular weight down to 42.24 kDa. This size difference only mildly improved the solubility of the fusion and thus suggests that other factors influence the stability. Issues concerning protein folding were initially explored by trying to denature and refold the proteins from a urea solution. Whilst this process did not solve the issues, the effect the solution's ionic strength had on the solubility of the protein points to the chemistry of the solutions as being a source of the problems. Consequently, buffer optimisation can be explored to find conditions suitable for *in vitro* assembly of the chimeric protein.

The truncated *Staphylococcus aureus* MR11 endolysin fused to DpsA yielded a construct with improved solubility. Furthermore, its assembly could be disrupted by the addition of urea. Whilst it was not carried out, it can be suggested that hetero-oligomeric assembly may be possible between DpsA and DpsAMR11t. This can be used to limit the number of functional groups within the dodecamer facilitating an improved solubility. Based on the DpsAmCh assembly displaying a weakened sensitivity to urea mediated dissociation (between 2 and 4 M urea) when compared to DpsA (between 4 and 5 M urea), a fewer number of functionalised subunits may even improve stability of the functionalised cage. Furthermore, the formation of a multi-functional cage, using DpsAmCh in combination with DpsAMR11t can allow the ability of the truncated MR11 domain to bind *S. aureus* cells to be tested by detecting the fluorescent signal.

A current hurdle for the preparation of chemically engineered protein shells or cage confined material synthesis is the requirement for harsh conditions; high temperatures and use of organic solvents for example, conditions which limit the usage of many protein based materials (Uchida et al., 2007). This study has shown how DpsC may provide a stable cage for

such procedures. Stability is not uncommon in the Dps family considering the characterisation of Dps from extreme environments (Franceschini et al., 2006), (Wiedenheft et al., 2005). Taking advantage of the growing understanding gained from biochemical and structural characterisation of novel Dps proteins, an approach highlighted by Ardejani et al., (2011) can be used which utilises *in silico* and *in vitro* techniques to engineer enhanced environment-specific stable protein cages. For instance the extended dimer interface of DpsC may be combined with an altered hydrophobic cavity displayed to increase thermal stability but maintain iron mineralisation (Swift et al., 2006). The alkylation of the cage surface by addition of long-chain primary amines to carboxylic groups would create an organic solvent soluble cage (Escosura et al., 2009). When combined, the enhanced thermostable, solvent soluble protein cage may be exploited for novel material catalysis or synthesis.

7.8 Future Work

The work presented in this thesis has provided an insight into the differences between the three *Sc*Dps. The work has also produced a number of intriguing questions relating to the structure and function of the *Sc*Dps proteins and their potential for development into bio-nanotechnological tools. Numerous studies can be carried out which can explore certain areas of this research in greater detail. A list of potential topics for future research follows;

- Accurate sizing of all the proteins explored in this study will provide a more in depth understanding of their oligomeric assembly in solution. Of particular interest are the deleted tail variants of DpsA (DpsADCT, DpsADNT and DpsADTM). This can be carried out by more accurate sizing techniques such as gel filtration or dynamic light scattering, although a multi technique approach would be beneficial.
- The assembly of DpsB can be briefly explored by manipulating the chemistry of the solution and even the sizing technique. Cross-linking may be an option for preserving the assembly if it is weakly associated.
- The DNA binding conditions of the *Sc*Dps can be explored in greater detail exploring multiple conditions which may include:
    - Inclusion of divalent metal ions
    - pH and ionic strength of reactions
    - Use of larger/smaller DNA fragments
    - Alternative DNA structure/topologies
- Examine the abilities of the proteins to mineralise alternative metals into a core

- Explore the stability of DpsC in other denaturing/extreme conditions using alternative assembly monitoring techniques such as dynamic light scattering or thermal shift assays
- Use PISA analysis and virtual alanine scanning results to infer key residues or hotspots for *in vitro* mutational studies in order to assess their effects on assembly/stability of the Dps dodecamers. Significant hotspots include;
    - DpsA; Trp-45, the highest ΔΔG (complex) value for the dimer interface (3.02 kcal/mol)
    - DpsA; Arg-76, the highest ΔΔG (complex) value for the "ferritin-like" interface (4.89 kcal/mol)
    - DpsA; Phe-52, the highest ΔΔG (complex) value for the "Dps-like" interface (2.65 kcal/mol)
    - DpsA; Trp-155, the only hotspot involved in two interfaces, (both trimer interfaces)
    - DpsC; Arg-118, the highest ΔΔG (complex) value for the dimer interface (4.82 kcal/mol) and is interestingly present on the BC helix
    - DpsC; Arg-179, the highest ΔΔG (complex) value for the "ferritin-like" interface (3.39 kcal/mol)
    - DpsC; Gln-83, the highest ΔΔG (complex) value for the "Dps-like" interface (3.14 kcal/mol)
    - DpsC; Leu-197, Val-198 and His-199 all hotspots located on the C-tail.
- Investigate the current MR11 lysin functionalization of DpsA using hetero-oligomeric assembly to multi-functionalise the dodecamer, combining cell recognition with fluorescence. Test non-assembling DpsA variants fused with truncated lysins.
- Consider other functional groups which have alternative targeting features
- Explore other methods of functionalization such as a "dock and lock" system or Cys-tagging, which allow post assembly addition of functional groups. Hetero-oligomeric assembly can manipulate ratios of docking recipient subunits.

Publications

FACEY, P. D., HITCHINGS, M. D., SAAVEDRA-GARCIA, P., FERNANDEZ-MARTINEZ, L., DYSON, P. J. & DEL SOL, R. 2009. Streptomyces coelicolor Dps-like proteins: differential dual roles in response to stress during vegetative growth and in nucleoid condensation during reproductive cell division. *Mol Microbiol,* 73, 1186-202.

FACEY, P. D., SEVCIKOVA, B., NOVAKOVA, R., HITCHINGS, M. D., CRACK, J. C., KORMANEC, J., DYSON, P. J. & DEL SOL, R. 2011. The dpsA gene of Streptomyces coelicolor: induction of expression from a single promoter in response to environmental stress or during development. *PLoS One,* 6, e25593.

FACEY, P. D., HITCHINGS, M. D., WILLIAMS, J. S., SKIBINSKI, D. O., DYSON, P. J. & DEL SOL, R. 2013. The evolution of an osmotically inducible dps in the genus Streptomyces. *PLoS One,* 8, e60772.

ABDUL-TEHRANI, H., HUDSON, A. J., CHANG, Y. S., TIMMS, A. R., HAWKINS, C., WILLIAMS, J. M., HARRISON, P. M., GUEST, J. R. & ANDREWS, S. C. 1999. Ferritin mutants of Escherichia coli are iron deficient and growth impaired, and fur mutants are iron deficient. *J Bacteriol,* 181, 1415-28.

ALALEONA, F., FRANCESCHINI, S., CECI, P., ILARI, A. & CHIANCONE, E. 2010. Thermosynechococcus elongatus DpsA binds Zn(II) at a unique three histidine-containing ferroxidase center and utilizes O2 as iron oxidant with very high efficiency, unlike the typical Dps proteins. *FEBS J,* 277, 903-17.

ALMIRÓN, M., LINK, A. J., FURLONG, D. & KOLTER, R. 1992. A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev,* 6, 2646-54.

ANDREWS, S. C. 2010. The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochim Biophys Acta,* 1800, 691-705.

ANDREWS, S. C., LE BRUN, N. E., BARYNIN, V., THOMSON, A. J., MOORE, G. R., GUEST, J. R. & HARRISON, P. M. 1995. Site-directed replacement of the coaxial heme ligands of bacterioferritin generates heme-free variants. *J Biol Chem,* 270, 23268-74.

ARDEJANI, M. S., CHOK, X. L., FOO, C. J. & ORNER, B. P. 2013. Complete shift of ferritin oligomerization toward nanocage assembly via engineered protein-protein interactions. *Chem Commun (Camb),* 49, 3528-30.

ARDEJANI, M. S., LI, N. X. & ORNER, B. P. 2011. Stabilization of a protein nanocage through the plugging of a protein-protein interfacial water pocket. *Biochemistry,* 50, 4029-37.

ARDINI, M., FIORILLO, A., FITTIPALDI, M., STEFANINI, S., GATTESCHI, D., ILARI, A. & CHIANCONE, E. 2013. Kineococcus radiotolerans Dps forms a heteronuclear Mn-Fe ferroxidase center that may explain the Mn-dependent protection against oxidative stress. *Biochim Biophys Acta,* 1830, 3745-55.

BAKKER, G. R. & BOYER, R. F. 1986. Iron incorporation into apoferritin. The role of apoferritin as a ferroxidase. *J Biol Chem,* 261, 13182-5.

BELLAPADRONA, G., ARDINI, M., CECI, P., STEFANINI, S. & CHIANCONE, E. 2010. Dps proteins prevent Fenton-mediated oxidative damage by trapping hydroxyl radicals within the protein shell. *Free Radic Biol Med,* 48, 292-7.

BELLAPADRONA, G., CHIARALUCE, R., CONSALVI, V., ILARI, A., STEFANINI, S. & CHIANCONE, E. 2007. The mutations Lys 114 --> Gln and Asp 126 --> Asn disrupt an intersubunit salt bridge and convert Listeria innocua Dps into its natural mutant Listeria monocytogenes Dps. Effects on protein stability at Low pH. *Proteins,* 66, 975-83.

BELLAPADRONA, G., STEFANINI, S., ZAMPARELLI, C., THEIL, E. C. & CHIANCONE, E. 2009. Iron translocation into and out of Listeria innocua Dps and size distribution of

the protein-enclosed nanomineral are modulated by the electrostatic gradient at the 3-fold "ferritin-like" pores. *J Biol Chem,* 284, 19101-9.

BENNETT, A. F. & LENSKI, R. E. 1997. Phenotypic and evolutionary adaptation of a model bacterial system to stressful thermal environments. *EXS,* 83, 135-54.

BENNION, B. J. & DAGGETT, V. 2003. The molecular basis for the chemical denaturation of proteins by urea. *Proc Natl Acad Sci U S A,* 100, 5142-7.

BENTLEY, S. D., CHATER, K. F., CERDENO-TARRAGA, A. M., CHALLIS, G. L., THOMSON, N. R., JAMES, K. D., HARRIS, D. E., QUAIL, M. A., KIESER, H., HARPER, D., BATEMAN, A., BROWN, S., CHANDRA, G., CHEN, C. W., COLLINS, M., CRONIN, A., FRASER, A., GOBLE, A., HIDALGO, J., HORNSBY, T., HOWARTH, S., HUANG, C. H., KIESER, T., LARKE, L., MURPHY, L., OLIVER, K., O'NEIL, S., RABBINOWITSCH, E., RAJANDREAM, M. A., RUTHERFORD, K., RUTTER, S., SEEGER, K., SAUNDERS, D., SHARP, S., SQUARES, R., SQUARES, S., TAYLOR, K., WARREN, T., WIETZORREK, A., WOODWARD, J., BARRELL, B. G., PARKHILL, J. & HOPWOOD, D. A. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature,* 417, 141-7.

BERTINI, I., LALLI, D., MANGANI, S., POZZI, C., ROSA, C., THEIL, E. C. & TURANO, P. 2012. Structural insights into the ferroxidase site of ferritins from higher eukaryotes. *J Am Chem Soc,* 134, 6169-76.

BHATTACHARYYA, G. & GROVE, A. 2007. The N-terminal extensions of Deinococcus radiodurans Dps-1 mediate DNA major groove interactions as well as assembly of the dodecamer. *J Biol Chem,* 282, 11921-30.

BOU-ABDALLAH, F. 2010. The iron redox and hydrolysis chemistry of the ferritins. *Biochim Biophys Acta,* 1800, 719-31.

BOU-ABDALLAH, F., AROSIO, P., SANTAMBROGIO, P., YANG, X., JANUS-CHANDLER, C. & CHASTEEN, N. D. 2002. Ferrous ion binding to recombinant human H-chain ferritin. An isothermal titration calorimetry study. *Biochemistry,* 41, 11184-91.

BOUGHAMMOURA, A., EXPERT, D. & FRANZA, T. 2012. Role of the Dickeya dadantii Dps protein. *Biometals,* 25, 423-33.

BRIAT, J. F. 1992. Iron assimilation and storage in prokaryotes. *J Gen Microbiol,* 138, 2475-83.

CARRONDO, M. A. 2003. Ferritins, iron uptake and storage from the bacterioferritin viewpoint. *EMBO J,* 22, 1959-68.

CECI, P., CELLAI, S., FALVO, E., RIVETTI, C., ROSSI, G. L. & CHIANCONE, E. 2004. DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res,* 32, 5935-44.

CECI, P., ILARI, A., FALVO, E. & CHIANCONE, E. 2003. The Dps protein of Agrobacterium tumefaciens does not bind to DNA but protects it toward oxidative cleavage: x-ray crystal structure, iron binding, and hydroxyl-radical scavenging properties. *J Biol Chem,* 278, 20319-26.

CECI, P., ILARI, A., FALVO, E., GIANGIACOMO, L. & CHIANCONE, E. 2005. Reassessment of protein stability, DNA binding, and protection of Mycobacterium smegmatis Dps. *J Biol Chem,* 280, 34776-85.

CECI, P., MANGIAROTTI, L., RIVETTI, C. & CHIANCONE, E. 2007. The neutrophil-activating Dps protein of Helicobacter pylori, HP-NAP, adopts a mechanism different from Escherichia coli Dps to bind and condense DNA. *Nucleic Acids Res,* 35, 2247-56.

CHAMPNESS, W. C. 1988. New loci required for Streptomyces coelicolor morphological and physiological differentiation. *J Bacteriol,* 170, 1168-74.

CHASTEEN, N. D. & HARRISON, P. M. 1999. Mineralization in ferritin: an efficient means of iron storage. *J Struct Biol,* 126, 182-94.

CHATER, K. F. 2006. Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos Trans R Soc Lond B Biol Sci,* 361, 761-8.

CHIANCONE, E. & CECI, P. 2010. The multifaceted capacity of Dps proteins to combat bacterial stress conditions: Detoxification of iron and hydrogen peroxide and DNA binding. *Biochim Biophys Acta,* 1800, 798-805.

CHIANCONE, E., CECI, P., ILARI, A., RIBACCHI, F. & STEFANINI, S. 2004. Iron and proteins for iron storage and detoxification. *Biometals,* 17, 197-202.

CHIARALUCE, R., CONSALVI, V., CAVALLO, S., ILARI, A., STEFANINI, S. & CHIANCONE, E. 2000. The unusual dodecameric ferritin from Listeria innocua dissociates below pH 2.0. *Eur J Biochem,* 267, 5733-41.

CHOLI-PAPADOPOULOU, T., KOTTAKIS, F., PAPADOPOULOS, G. & PENDAS, S. 2011. Helicobacter pylori neutrophil activating protein as target for new drugs against H. pylori inflammation. *World J Gastroenterol,* 17, 2585-91.

CHOWDHURY, R. P., VIJAYABASKAR, M. S., VISHVESHWARA, S. & CHATTERJI, D. 2008. Molecular mechanism of in vitro oligomerization of Dps from Mycobacterium smegmatis: mutations of the residues identified by "interface cluster" analysis. *Biochemistry,* 47, 11110-7.

CLAESSEN, D., DE JONG, W., DIJKHUIZEN, L. & WOSTEN, H. A. 2006. Regulation of Streptomyces development: reach for the sky! *Trends Microbiol,* 14, 313-9.

COLBURN-CLIFFORD, J. M., SCHERF, J. M. & ALLEN, C. 2010. Ralstonia solanacearum Dps contributes to oxidative stress tolerance and to colonization of and virulence on tomato plants. *Appl Environ Microbiol,* 76, 7392-9.

COLOVOS, C. & YEATES, T. O. 1993. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci,* 2, 1511-9.

COTRUVO, J. A. & STUBBE, J. 2012. Metallation and mismetallation of iron and manganese proteins in vitro and in vivo: the class I ribonucleotide reductases as a case study. *Metallomics,* 4, 1020-36.

CRICHTON, R. R. & DECLERCQ, J. P. 2010. X-ray structures of ferritins and related proteins. *Biochim Biophys Acta,* 1800, 706-18.

CROW, A., LAWSON, T. L., LEWIN, A., MOORE, G. R. & LE BRUN, N. E. 2009. Structural basis for iron mineralization by bacterioferritin. *J Am Chem Soc,* 131, 6808-13.

CUYPERS, M. G., MITCHELL, E. P., ROMÃO, C. V. & MCSWEENEY, S. M. 2007. The crystal structure of the Dps2 from Deinococcus radiodurans reveals an unusual pore profile with a non-specific metal binding site. *J Mol Biol,* 371, 787-99.

D'AGOSTINO, D., RACIOPPI, M., FILIANOTI, A., DI GIANFRANCESCO, L., CODOLO, G., FASSAN, M., MUNARI, F., RUGGE, M., D'ELIOS, M. M., DE BERNARD, M., PAGANO, F. & BASSI, P. 2012. [Therapy for non-muscle invasive bladder cancer: HP-NAP]. *Urologia,* 79, 142-8.

D'ELIOS, M. M., AMEDEI, A., CAPPON, A., DEL PRETE, G. & DE BERNARD, M. 2007. The neutrophil-activating protein of Helicobacter pylori (HP-NAP) as an immune modulating agent. *FEMS Immunol Med Microbiol,* 50, 157-64.

DAVIES, J. & DAVIES, D. 2010. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev,* 74, 417-33.

DE BERNARD, M. & D'ELIOS, M. M. 2010. The immune modulating activity of the Helicobacter pylori HP-NAP: Friend or foe? *Toxicon,* 56, 1186-92.

DE VAL, N., DECLERCQ, J. P., LIM, C. K. & CRICHTON, R. R. 2012. Structural analysis of haemin demetallation by L-chain apoferritins. *J Inorg Biochem,* 112, 77-84.

DEDMAN, D. J., TREFFRY, A., CANDY, J. M., TAYLOR, G. A., MORRIS, C. M., BLOXHAM, C. A., PERRY, R. H., EDWARDSON, J. A. & HARRISON, P. M. 1992. Iron and aluminium in relation to brain ferritin in normal individuals and Alzheimer's-disease and chronic renal-dialysis patients. *Biochem J,* 287 (Pt 2), 509-14.

DOLINSKY, T. J., CZODROWSKI, P., LI, H., NIELSEN, J. E., JENSEN, J. H., KLEBE, G. & BAKER, N. A. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res,* 35, W522-5.

DOLL, T. A., RAMAN, S., DEY, R. & BURKHARD, P. 2013. Nanoscale assemblies and their biomedical applications. *J R Soc Interface,* 10, 20120740.

ELLIOT, M. A., KAROONUTHAISIRI, N., HUANG, J., BIBB, M. J., COHEN, S. N., KAO, C. M. & BUTTNER, M. J. 2003. The chaplins: a family of hydrophobic cell-surface

proteins involved in aerial mycelium formation in Streptomyces coelicolor. *Genes Dev,* 17, 1727-40.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66, 486-501.

ENSIGN, D., YOUNG, M. & DOUGLAS, T. 2004. Photocatalytic synthesis of copper colloids from CuII by the ferrihydrite core of ferritin. *Inorg Chem,* 43, 3441-6.

ESCOSURA, A. D. L., NOLTE, R. J. M. & CORNELISSEN, J. J. L. M. 2009. Viruses and protein cages as nanocontainers and nanoreactors.

EYAL, E., GERZON, S., POTAPOV, V., EDELMAN, M. & SOBOLEV, V. 2005. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J Mol Biol,* 351, 431-42.

FACEY, P. D., HITCHINGS, M. D., SAAVEDRA-GARCIA, P., FERNANDEZ-MARTINEZ, L., DYSON, P. J. & DEL SOL, R. 2009. Streptomyces coelicolor Dps-like proteins: differential dual roles in response to stress during vegetative growth and in nucleoid condensation during reproductive cell division. *Mol Microbiol,* 73, 1186-202.

FACEY, P. D., HITCHINGS, M. D., WILLIAMS, J. S., SKIBINSKI, D. O., DYSON, P. J. & DEL SOL, R. 2013. The evolution of an osmotically inducible dps in the genus Streptomyces. *PLoS One,* 8, e60772.

FACEY, P. D., SEVCIKOVA, B., NOVAKOVA, R., HITCHINGS, M. D., CRACK, J. C., KORMANEC, J., DYSON, P. J. & DEL SOL, R. 2011. The dpsA gene of Streptomyces coelicolor: induction of expression from a single promoter in response to environmental stress or during development. *PLoS One,* 6, e25593.

FAN, R., BOYLE, A. L., CHEONG, V. V., NG, S. L. & ORNER, B. P. 2009. A helix swapping study of two protein cages. *Biochemistry,* 48, 5623-30.

FISCHETTI, V. A. 2010. Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *Int J Med Microbiol,* 300, 357-62.

FLARDH, K., FINDLAY, K. C. & CHATER, K. F. 1999. Association of early sporulation genes with suggested developmental decision points in Streptomyces coelicolor A3(2). *Microbiology,* 145 (Pt 9), 2229-43.

FLENNIKEN, M. L., UCHIDA, M., LIEPOLD, L. O., KANG, S., YOUNG, M. J. & DOUGLAS, T. 2009. A library of protein cage architectures as nanomaterials. *Curr Top Microbiol Immunol,* 327, 71-93.

FORD, G. C., HARRISON, P. M., RICE, D. W., SMITH, J. M., TREFFRY, A., WHITE, J. L. & YARIV, J. 1984. Ferritin: design and formation of an iron-storage molecule. *Philos Trans R Soc Lond B Biol Sci,* 304, 551-65.

FRANCESCHINI, S., CECI, P., ALALEONA, F., CHIANCONE, E. & ILARI, A. 2006. Antioxidant Dps protein from the thermophilic cyanobacterium Thermosynechococcus elongatus. *FEBS J,* 273, 4913-28.

FRENKIEL-KRISPIN, D., LEVIN-ZAIDMAN, S., SHIMONI, E., WOLF, S. G., WACHTEL, E. J., ARAD, T., FINKEL, S. E., KOLTER, R. & MINSKY, A. 2001. Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection. *Embo j,* 20, 1184-91.

FROLOW, F., KALB, A. J. & YARIV, J. 1994. Structure of a unique twofold symmetric haem-binding site. *Nat Struct Biol,* 1, 453-60.

GAUSS, G. H., BENAS, P., WIEDENHEFT, B., YOUNG, M., DOUGLAS, T. & LAWRENCE, C. M. 2006. Structure of the DPS-like protein from Sulfolobus solfataricus reveals a bacterioferritin-like dimetal binding site within a DPS-like dodecameric assembly. *Biochemistry,* 45, 10815-27.

GENINATTI CRICH, S., CUTRIN, J. C., LANZARDO, S., CONTI, L., KALMAN, F. K., SZABO, I., LAGO, N. R., IOLASCON, A. & AIME, S. 2012. Mn-loaded apoferritin: a highly sensitive MRI imaging probe for the detection and characterization of hepatocarcinoma lesions in a transgenic mouse model. *Contrast Media Mol Imaging,* 7, 281-8.

GERL, M. & JAENICKE, R. 1987. Mechanism of the self-assembly of apoferritin from horse spleen. Cross-linking and spectroscopic analysis. *Eur Biophys J,* 15, 103-9.

GONZÁLEZ-FLECHA, B. & DEMPLE, B. 1995. Metabolic sources of hydrogen peroxide in aerobically growing Escherichia coli. *J Biol Chem,* 270, 13681-7.

GRANIER, T., LANGLOIS D'ESTAINTOT, B., GALLOIS, B., CHEVALIER, J. M., PRÉCIGOUX, G., SANTAMBROGIO, P. & AROSIO, P. 2003. Structural description of the active sites of mouse L-chain ferritin at 1.2 A resolution. *J Biol Inorg Chem,* 8, 105-11.

GRANT, R. A., FILMAN, D. J., FINKEL, S. E., KOLTER, R. & HOGLE, J. M. 1998. The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol,* 5, 294-303.

GRAY, D. I., GOODAY, G. W. & PROSSER, J. I. 1990. Apical hyphal extension in Streptomyces coelicolor A3(2). *J Gen Microbiol,* 136, 1077-84.

GROVE, A. & WILKINSON, S. P. 2005. Differential DNA binding and protection by dimeric and dodecameric forms of the ferritin homolog Dps from Deinococcus radiodurans. *J Mol Biol,* 347, 495-508.

GROVES, M. & LUCANA, D. O. 2010. *Adaptation to oxidative stress by Gram-positive bacteria: the redoxsensing system HbpS-SenS-SenR from Streptomyces reticuli,* Spain, Formatex.

GUPTA, A. K. & GUPTA, M. 2005. Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. *Biomaterials,* 26, 3995-4021.

GUPTA, S. & CHATTERJI, D. 2003. Bimodal protection of DNA by Mycobacterium smegmatis DNA-binding protein from stationary phase cells. *J Biol Chem,* 278, 5235-41.

GUPTA, V., GUPTA, R. K., KHARE, G., SALUNKE, D. M. & TYAGI, A. K. 2009. Crystal structure of Bfr A from Mycobacterium tuberculosis: incorporation of selenomethionine results in cleavage and demetallation of haem. *PLoS One,* 4, e8028.

HA, Y., SHI, D., SMALL, G. W., THEIL, E. C. & ALLEWELL, N. M. 1999. Crystal structure of bullfrog M ferritin at 2.8 A resolution: analysis of subunit interactions and the binuclear metal center. *J Biol Inorg Chem,* 4, 243-56.

HAIKARAINEN, T. & PAPAGEORGIOU, A. C. 2010. Dps-like proteins: structural and functional insights into a versatile protein family. *Cell Mol Life Sci,* 67, 341-51.

HAIKARAINEN, T., PATURI, P., LINDEN, J., HAATAJA, S., MEYER-KLAUCKE, W., FINNE, J. & PAPAGEORGIOU, A. C. 2011. Magnetic properties and structural characterization of iron oxide nanoparticles formed by Streptococcus suis Dpr and four mutants. *J Biol Inorg Chem,* 16, 799-807.

HALLIWELL, B. & GUTTERIDGE, J. M. 1984. Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem J,* 219, 1-14.

HALSEY, T. A., VAZQUEZ-TORRES, A., GRAVDAHL, D. J., FANG, F. C. & LIBBY, S. J. 2004. The ferritin-like Dps protein is required for Salmonella enterica serovar Typhimurium oxidative stress resistance and virulence. *Infect Immun,* 72, 1155-8.

HARRISON, P. M. & AROSIO, P. 1996. The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim Biophys Acta,* 1275, 161-203.

HAUN, J. B., YOON, T. J., LEE, H. & WEISSLEDER, R. 2010. Magnetic nanoparticle biosensors. *Wiley Interdiscip Rev Nanomed Nanobiotechnol,* 2, 291-304.

HEMPSTEAD, P. D., YEWDALL, S. J., FERNIE, A. R., LAWSON, D. M., ARTYMIUK, P. J., RICE, D. W., FORD, G. C. & HARRISON, P. M. 1997. Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. *J Mol Biol,* 268, 424-48.

HUDSON, A. J., ANDREWS, S. C., HAWKINS, C., WILLIAMS, J. M., IZUHARA, M., MELDRUM, F. C., MANN, S., HARRISON, P. M. & GUEST, J. R. 1993. Overproduction, purification and characterization of the Escherichia coli ferritin. *Eur J Biochem,* 218, 985-95.

HUERGO, L. F., RAHMAN, H., IBRAHIMOVIC, A., DAY, C. J. & KOROLIK, V. 2013. Campylobacter jejuni Dps protein binds DNA in the presence of iron or hydrogen peroxide. *J Bacteriol,* 195, 1970-8.

HWANG, J., KREBS, C., HUYNH, B. H., EDMONDSON, D. E., THEIL, E. C. & PENNER-HAHN, J. E. 2000. A short Fe-Fe distance in peroxodiferric ferritin: control of Fe substrate versus cofactor decay? *Science,* 287, 122-5.

ILARI, A., CECI, P., FERRARI, D., ROSSI, G. L. & CHIANCONE, E. 2002. Iron incorporation into Escherichia coli Dps gives rise to a ferritin-like microcrystalline core. *J Biol Chem,* 277, 37619-23.

ILARI, A., LATELLA, M. C., CECI, P., RIBACCHI, F., SU, M., GIANGIACOMO, L., STEFANINI, S., CHASTEEN, N. D. & CHIANCONE, E. 2005. The unusual intersubunit ferroxidase center of Listeria innocua Dps is required for hydrogen peroxide detoxification but not for iron uptake. A study with site-specific mutants. *Biochemistry,* 44, 5579-87.

ILARI, A., STEFANINI, S., CHIANCONE, E. & TSERNOGLOU, D. 2000. The dodecameric ferritin from Listeria innocua contains a novel intersubunit iron-binding site. *Nat Struct Biol,* 7, 38-43.

ILBERT, M. & BONNEFOY, V. 2013. Insight into the evolution of the iron oxidation pathways. *Biochim Biophys Acta,* 1827, 161-75.

ISHIKAWA, K., YAMADA, K., KUMAGAI, S., SANO, K.-I., SHIBA, K., YAMASHITA, I. & KOBAYASHI, M. 2008. Adsorption Properties of a Gold-Binding Peptide Assessed by its Attachment to a Recombinant Apoferritin Molecule. *Applied Physics Express 1 (2008).*

ISHIKAWA, T., MIZUNOE, Y., KAWABATA, S., TAKADE, A., HARADA, M., WAI, S. N. & YOSHIDA, S. 2003. The iron-binding protein Dps confers hydrogen peroxide stress resistance to Campylobacter jejuni. *J Bacteriol,* 185, 1010-7.

JANOWSKI, R., AUERBACH-NEVO, T. & WEISS, M. S. 2008. Bacterioferritin from Mycobacterium smegmatis contains zinc in its di-nuclear site. *Protein Sci,* 17, 1138-50.

JEONG, K. C., HUNG, K. F., BAUMLER, D. J., BYRD, J. J. & KASPAR, C. W. 2008. Acid stress damage of DNA is prevented by Dps binding in Escherichia coli O157:H7. *BMC Microbiol,* 8, 181.

KABSCH, W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr,* 66, 125-32.

KANG, S., SUCI, P. A., BROOMELL, C. C., IWAHORI, K., KOBAYASHI, M., YAMASHITA, I., YOUNG, M. & DOUGLAS, T. 2009. Janus-like protein cages. Spatially controlled dual-functional surface modifications of protein cages. *Nano Lett,* 9, 2360-6.

KAPPLER, A. & STRAUB, K. 2005. Geomicrobiological cycling of iron. *Molecular Geomicrobiology,* 59, 85-108.

KHARE, G., GUPTA, V., NANGPAL, P., GUPTA, R. K., SAUTER, N. K. & TYAGI, A. K. 2011. Ferritin structure from Mycobacterium tuberculosis: comparative study with

homologues identifies extended C-terminus involved in ferroxidase activity. *PLoS One,* 6, e18570.

KIM, S. E., AHN, K. Y., PARK, J. S., KIM, K. R., LEE, K. E., HAN, S. S. & LEE, J. 2011. Fluorescent ferritin nanoparticles and application to the aptamer sensor. *Anal Chem,* 83, 5834-43.

KO, Y., KIM, Y., BAEK, H. & CHO, J. 2011. Electrically bistable properties of layer-by-layer assembled multilayers based on protein nanoparticles. *ACS Nano,* 5, 9918-26.

KORTEMME, T. & BAKER, D. 2002. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci U S A,* 99, 14116-21.

KORTEMME, T., KIM, D. E. & BAKER, D. 2004. Computational alanine scanning of protein-protein interfaces. *Sci STKE,* 2004, pl2.

KRAMER, R. M., LI, C., CARTER, D. C., STONE, M. O. & NAIK, R. R. 2004. Engineered protein cages for nanomaterial synthesis. *J Am Chem Soc,* 126, 13282-6.

KRETZER, J. W., LEHMANN, R., SCHMELCHER, M., BANZ, M., KIM, K. P., KORN, C. & LOESSNER, M. J. 2007. Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. *Appl Environ Microbiol,* 73, 1992-2000.

KRISSINEL, E. & HENRICK, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol,* 372, 774-97.

KUMAR, S. & NUSSINOV, R. 2002. Close-range electrostatic interactions in proteins. *Chembiochem,* 3, 604-17.

KYTE, J. & DOOLITTLE, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol,* 157, 105-32.

LAI, C.-H., HSU, Y.-M., WANG, H.-J. & WANG, W.-C. 2013. Manipulation of host cholesterol by Helicobacter pylori for their beneficial ecological niche. *BioMedicine,* 3, 27-33.

LAWSON, D. M., ARTYMIUK, P. J., YEWDALL, S. J., SMITH, J. M., LIVINGSTONE, J. C., TREFFRY, A., LUZZAGO, A., LEVI, S., AROSIO, P. & CESARENI, G. 1991. Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature,* 349, 541-4.

LE BRUN, N. E., CROW, A., MURPHY, M. E., MAUK, A. G. & MOORE, G. R. 2010. Iron core mineralisation in prokaryotic ferritins. *Biochim Biophys Acta,* 1800, 732-44.

LEONG, L. M., TAN, B. H. & HO, K. K. 1992. A specific stain for the detection of nonheme iron proteins in polyacrylamide gels. *Anal Biochem,* 207, 317-20.

LEVI, S., SALFELD, J., FRANCESCHINELLI, F., COZZI, A., DORNER, M. H. & AROSIO, P. 1989. Expression and structural and functional properties of human ferritin L-chain from Escherichia coli. *Biochemistry,* 28, 5179-84.

LI, K., ZHANG, Z.-P., LUO, M., YU, X., HAN, Y., WEI, H.-P., CUI, Z.-Q. & ZHANG, X.-E. 2011. Multifunctional ferritin cage nanostructures for fluorescence and MR imaging of tumor cells.

LINDER, M. C., KAKAVANDI, H. R., MILLER, P., WIRTH, P. L. & NAGEL, G. M. 1989. Dissociation of ferritins. *Arch Biochem Biophys,* 269, 485-96.

LIOLIOS, K., SIKORSKI, J., JANDO, M., LAPIDUS, A., COPELAND, A., GLAVINA, T., DEL, R., NOLAN, M., LUCAS, S., TICE, H., CHENG, J. F., HAN, C., WOYKE, T., GOODWIN, L., PITLUCK, S., IVANOVA, N., MAVROMATIS, K., MIKHAILOVA, N., CHERTKOV, O., KUSKE, C., CHEN, A., PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y. J., JEFFRIES, C. D., DETTER, J. C., BRETTIN, T., ROHDE, M., GOKER, M., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ, P., KLENK, H. P. & KYRPIDES, N. C. 2010. Complete genome sequence of Thermobispora bispora type strain (R51). *Stand Genomic Sci,* 2, 318-26.

LIU, X., KIM, K., LEIGHTON, T. & THEIL, E. C. 2006. Paired Bacillus anthracis Dps (mini-ferritin) have different reactivities with peroxide. *J Biol Chem,* 281, 27827-35.

LOESSNER, M. J. 2005. Bacteriophage endolysins--current state of research and applications. *Curr Opin Microbiol,* 8, 480-7.

LOVELL, S. C., DAVIS, I. W., ARENDALL, W. B., DE BAKKER, P. I., WORD, J. M., PRISANT, M. G., RICHARDSON, J. S. & RICHARDSON, D. C. 2003. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins,* 50, 437-50.

LOW, L. Y., YANG, C., PEREGO, M., OSTERMAN, A. & LIDDINGTON, R. C. 2005. Structure and lytic activity of a Bacillus anthracis prophage endolysin. *J Biol Chem,* 280, 35433-9.

LUO, Y., HAN, Z., CHIN, S. M. & LINN, S. 1994. Three chemically distinct types of oxidants formed by iron-mediated Fenton reactions in the presence of DNA. *Proc Natl Acad Sci U S A,* 91, 12438-42.

MACEDO, S., ROMÃO, C. V., MITCHELL, E., MATIAS, P. M., LIU, M. Y., XAVIER, A. V., LEGALL, J., TEIXEIRA, M., LINDLEY, P. & CARRONDO, M. A. 2003. The nature of the di-iron site in the bacterioferritin from Desulfovibrio desulfuricans. *Nat Struct Biol,* 10, 285-90.

MAILLOUX, R. J. & HARPER, M. E. 2011. Uncoupling proteins and the control of mitochondrial reactive oxygen species production. *Free Radic Biol Med,* 51, 1106-15.

MATIAS, P. M., TATUR, J., CARRONDO, M. A. & HAGEN, W. R. 2005. Crystallization and preliminary X-ray characterization of a ferritin from the hyperthermophilic archaeon and anaerobe Pyrococcus furiosus. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 61, 503-6.

MATSUDA, A., FURUKAWA, K., SUZUKI, H., KAN, H., TSURUTA, H., MATSUMOTO, S., SHINJI, S. & TAJIRI, T. 2007. Does impaired TH1/TH2 balance cause postoperative infectious complications in colorectal cancer surgery? *J Surg Res,* 139, 15-21.

MATSUZAKI, S., RASHEL, M., UCHIYAMA, J., SAKURAI, S., UJIHARA, T., KURODA, M., IKEUCHI, M., TANI, T., FUJIEDA, M., WAKIGUCHI, H. & IMAI, S. 2005. Bacteriophage therapy: a revitalized therapy against bacterial infectious diseases. *J Infect Chemother,* 11, 211-9.

MAYER, M. J., GAREFALAKI, V., SPOERL, R., NARBAD, A. & MEIJERS, R. 2011. Structure-based modification of a Clostridium difficile-targeting endolysin affects activity and host range. *J Bacteriol,* 193, 5477-86.

MAYER, M. J., NARBAD, A. & GASSON, M. J. 2008. Molecular characterization of a Clostridium difficile bacteriophage and its cloned biologically active endolysin. *J Bacteriol,* 190, 6734-40.

MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography,* 40, 658-674.

MICHEL, F. M., HOSEIN, H. A., HAUSNER, D. B., DEBNATH, S., PARISE, J. B. & STRONGIN, D. R. 2010. Reactivity of ferritin and the structure of ferritin-derived ferrihydrite. *Biochim Biophys Acta,* 1800, 871-85.

MIYAMOTO, T., ASAHINA, Y., MIYAZAKI, S., SHIMIZU, H., OHTO, U., NOGUCHI, S. & SATOW, Y. 2011a. Structures of the SEp22 dodecamer, a Dps-like protein from Salmonella enterica subsp. enterica serovar Enteritidis. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 67, 17-22.

MIYAMOTO, T., ASAHINA, Y., MIYAZAKI, S., SHIMIZU, H., OHTO, U., NOGUCHI, S. & SATOW, Y. 2011b. Structures of the SEp22 dodecamer, a Dps-like protein from Salmonella enterica subsp. enterica serovar Enteritidis. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 67, 17-22.

MONAGHAN, T., BOSWELL, T. & MAHIDA, Y. R. 2008. Recent advances in Clostridium difficile-associated disease. *Gut,* 57, 850-60.

MOORE, G. R., KADIR, F. H., AL-MASSAD, F. K., LE BRUN, N. E., THOMSON, A. J., GREENWOOD, C., KEEN, J. N. & FINDLAY, J. B. 1994. Structural heterogeneity of Pseudomonas aeruginosa bacterioferritin. *Biochem J,* 304 (Pt 2), 493-7.

MUNRO, H. N., AZIZ, N., LEIBOLD, E. A., MURRAY, M., ROGERS, J., VASS, J. K. & WHITE, K. 1988. The ferritin genes: structure, expression, and regulation. *Ann N Y Acad Sci,* 526, 113-23.

NGUYEN, K. H., SMITH, L. T., XIAO, L., BHATTACHARYYA, G. & GROVE, A. 2012. On the stoichiometry of Deinococcus radiodurans Dps-1 binding to duplex DNA. *Proteins,* 80, 713-21.

NICODÈME, M., PERRIN, C., HOLS, P., BRACQUART, P. & GAILLARD, J. L. 2004. Identification of an iron-binding protein of the Dps family expressed by Streptococcus thermophilus. *Curr Microbiol,* **48,** 51-6.

OKUDA, M., IWAHORI, K., YAMASHITA, I. & YOSHIMURA, H. 2003. Fabrication of nickel and chromium nanoparticles using the protein cage of apoferritin. *Biotechnol Bioeng,* **84,** 187-94.

OU, H. Y., GUO, F. B. & ZHANG, C. T. 2003. Analysis of nucleotide distribution in the genome of Streptomyces coelicolor A3(2) using the Z curve method. *FEBS Lett,* **540,** 188-94.

PAN, Y. H., SADER, K., POWELL, J. J., BLELOCH, A., GASS, M., TRINICK, J., WARLEY, A., LI, A., BRYDSON, R. & BROWN, A. 2009. 3D morphology of the human hepatic ferritin mineral core: new evidence for a subunit structure revealed by single particle analysis of HAADF-STEM images. *J Struct Biol,* **166,** 22-31.

PAPINUTTO, E., DUNDON, W. G., PITULIS, N., BATTISTUTTA, R., MONTECUCCO, C. & ZANOTTI, G. 2002. Structure of two iron-binding proteins from Bacillus anthracis. *J Biol Chem,* **277,** 15093-8.

PARADKAR, A., TREFZER, A., CHAKRABURTTY, R. & STASSI, D. 2003. Streptomyces genetics: a genomic perspective. *Crit Rev Biotechnol,* **23,** 1-27.

PESEK, J., BUCHLER, R., ALBRECHT, R., BOLAND, W. & ZETH, K. 2011. Structure and mechanism of iron translocation by a Dps protein from Microbacterium arborescens. *J Biol Chem,* **286,** 34872-82.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* **25,** 1605-12.

PRADEL, N., SANTINI, C. L., BERNADAC, A., SHIH, Y. L., GOLDBERG, M. B. & WU, L. F. 2007. Polar positional information in Escherichia coli spherical cells. *Biochem Biophys Res Commun,* **353,** 493-500.

RAJASEKARAN, M. B., NILAPWAR, S., ANDREWS, S. C. & WATSON, K. A. 2010. EfeO-cupredoxins: major new members of the cupredoxin superfamily with roles in bacterial iron transport. *Biometals,* **23,** 1-17.

RAMACHANDRAN, M., YU, D., WANDERS, A., ESSAND, M. & ERIKSSON, F. 2013. An infection-enhanced oncolytic adenovirus secreting H. pylori neutrophil-activating protein with therapeutic effects on neuroendocrine tumors. *Mol Ther.*

RASHEL, M., UCHIYAMA, J., UJIHARA, T., UEHARA, Y., KURAMOTO, S., SUGIHARA, S., YAGYU, K., MURAOKA, A., SUGAI, M., HIRAMATSU, K., HONKE, K. & MATSUZAKI, S. 2007. Efficient elimination of multidrug-resistant Staphylococcus aureus by cloned lysin derived from bacteriophage phi MR11. *J Infect Dis,* **196,** 1237-47.

REN, B., TIBBELIN, G., KAJINO, T., ASAMI, O. & LADENSTEIN, R. 2003. The multi-layered structure of Dps with a novel di-nuclear ferroxidase center. *J Mol Biol,* 329, 467-77.

REON, B. J., NGUYEN, K. H., BHATTACHARYYA, G. & GROVE, A. 2012. Functional comparison of Deinococcus radiodurans Dps proteins suggests distinct in vivo roles. *Biochem J,* 447, 381-91.

ROMAGNANI, S. 2003. Lymphokine Production by Human T Cells in Disease States. http://dx.doi.org/10.1146/annurev.iy.12.040194.001303.

ROMÃO, C. V., MITCHELL, E. P. & MCSWEENEY, S. 2006. The crystal structure of Deinococcus radiodurans Dps protein (DR2263) reveals the presence of a novel metal centre in the N terminus. *J Biol Inorg Chem,* 11, 891-902.

ROY, S., GUPTA, S., DAS, S., SEKAR, K., CHATTERJI, D. & VIJAYAN, M. 2004. X-ray analysis of Mycobacterium smegmatis Dps and a comparative study involving other Dps and Dps-like molecules. *J Mol Biol,* 339, 1103-13.

ROY, S., SARASWATHI, R., CHATTERJI, D. & VIJAYAN, M. 2008. Structural studies on the second Mycobacterium smegmatis Dps: invariant and variable features of structure, assembly and function. *J Mol Biol,* 375, 948-59.

ROY, S., SARASWATHI, R., GUPTA, S., SEKAR, K., CHATTERJI, D. & VIJAYAN, M. 2007. Role of N and C-terminal tails in DNA binding and assembly in Dps: structural studies of Mycobacterium smegmatis Dps deletion mutants. *J Mol Biol,* 370, 752-67.

RUBAN-OSMIALOWSKA, B., JAKIMOWICZ, D., SMULCZYK-KRAWCZYSZYN, A., CHATER, K. F. & ZAKRZEWSKA-CZERWINSKA, J. 2006. Replisome localization in vegetative and aerial hyphae of Streptomyces coelicolor. *J Bacteriol,* 188, 7311-6.

SARASWATHI, R., PAIT CHOWDHURY, R., WILLIAMS, S. M., GHATAK, P. & CHATTERJI, D. 2009. The mycobacterial MsDps2 protein is a nucleoid-forming DNA binding protein regulated by sigma factors sigma and sigma. *PLoS One,* 4, e8017.

SCHWARTZ, J. K., LIU, X. S., TOSHA, T., DIEBOLD, A., THEIL, E. C. & SOLOMON, E. I. 2010. CD and MCD spectroscopic studies of the two Dps miniferritin proteins from Bacillus anthracis: role of O2 and H2O2 substrates in reactivity of the diiron catalytic centers. *Biochemistry,* 49, 10516-25.

SHANER, N. C., CAMPBELL, R. E., STEINBACH, P. A., GIEPMANS, B. N., PALMER, A. E. & TSIEN, R. Y. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol,* 22, 1567-72.

SHANNON, R. 1976. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallographica Section A,* 32, 751-767.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SÖDING, J., THOMPSON, J. D. & HIGGINS, D.

G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7, 539.

SKUBÁK, P., MURSHUDOV, G. N. & PANNU, N. S. 2004. Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr D Biol Crystallogr,* 60, 2196-201.

STEFANINI, S., VECCHINI, P. & CHIANCONE, E. 1987. On the mechanism of horse spleen apoferritin assembly: a sedimentation velocity and circular dichroism study. *Biochemistry,* 26, 1831-7.

STEIN, N. 2008. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography,* 41, 641-643.

STEINGRUBE, V. A., WILSON, R. W., BROWN, B. A., JOST, K. C., JR., BLACKLOCK, Z., GIBSON, J. L. & WALLACE, R. J., JR. 1997. Rapid identification of clinically significant species and taxa of aerobic actinomycetes, including Actinomadura, Gordona, Nocardia, Rhodococcus, Streptomyces, and Tsukamurella isolates, by DNA amplification and restriction endonuclease analysis. *J Clin Microbiol,* 35, 817-22.

STILLMAN, T. J., HEMPSTEAD, P. D., ARTYMIUK, P. J., ANDREWS, S. C., HUDSON, A. J., TREFFRY, A., GUEST, J. R. & HARRISON, P. M. 2001. The high-resolution X-ray crystallographic structure of the ferritin (EcFtnA) of Escherichia coli; comparison with human H ferritin (HuHF) and the structures of the Fe(3+) and Zn(2+) derivatives. *J Mol Biol,* 307, 587-603.

STILLMAN, T. J., UPADHYAY, M., NORTE, V. A., SEDELNIKOVA, S. E., CARRADUS, M., TZOKOV, S., BULLOUGH, P. A., SHEARMAN, C. A., GASSON, M. J., WILLIAMS, C. H., ARTYMIUK, P. J. & GREEN, J. 2005. The crystal structures of Lactococcus lactis MG1363 Dps proteins reveal the presence of an N-terminal helix that is required for DNA binding. *Mol Microbiol,* 57, 1101-12.

SU, M., CAVALLO, S., STEFANINI, S., CHIANCONE, E. & CHASTEEN, N. D. 2005. The so-called Listeria innocua ferritin is a Dps protein. Iron incorporation, detoxification, and DNA protection properties. *Biochemistry,* 44, 5572-8.

SULAKVELIDZE, A., ALAVIDZE, Z. & MORRIS, J. G. 2001. Bacteriophage therapy. *Antimicrob Agents Chemother,* 45, 649-59.

SUN, C., YANG, H., YUAN, Y., TIAN, X., WANG, L., GUO, Y., XU, L., LEI, J., GAO, N., ANDERSON, G. J., LIANG, X. J., CHEN, C., ZHAO, Y. & NIE, G. 2011. Controlling assembly of paired gold clusters within apoferritin nanoreactor for in vivo kidney targeting and biomedical imaging. *J Am Chem Soc,* 133, 8617-24.

SWARTZ, L., KUCHINSKAS, M., LI, H., POULOS, T. L. & LANZILOTTA, W. N. 2006. Redox-dependent structural changes in the Azotobacter vinelandii bacterioferritin: new insights into the ferroxidase and iron transport mechanism. *Biochemistry,* 45, 4421-8.

SWIFT, J., WEHBI, W. A., KELLY, B. D., STOWELL, X. F., SAVEN, J. G. & DMOCHOWSKI, I. J. 2006. Design of functional ferritin-like proteins with hydrophobic cavities. *J Am Chem Soc,* 128, 6611-9.

TAKAHASHI, T. & KUYUCAK, S. 2003. Functional properties of threefold and fourfold channels in ferritin deduced from electrostatic calculations. *Biophys J,* 84, 2256-63.

TATUR, J., HAGEN, W. R. & MATIAS, P. M. 2007. Crystal structure of the ferritin from the hyperthermophilic archaeal anaerobe Pyrococcus furiosus. *J Biol Inorg Chem,* 12, 615-30.

THEORET, J. R., COOPER, K. K., ZEKARIAS, B., ROLAND, K. L., LAW, B. F., CURTISS, R. & JOENS, L. A. 2012. The Campylobacter jejuni Dps homologue is important for in vitro biofilm formation and cecal colonization of poultry and may serve as a protective antigen for vaccination. *Clin Vaccine Immunol,* 19, 1426-31.

THUMIGER, A., POLENGHI, A., PAPINUTTO, E., BATTISTUTTA, R., MONTECUCCO, C. & ZANOTTI, G. 2006. Crystal structure of antigen TpF1 from Treponema pallidum. *Proteins,* 62, 827-30.

TREFFRY, A., BAUMINGER, E. R., HECHEL, D., HODSON, N. W., NOWIK, I., YEWDALL, S. J. & HARRISON, P. M. 1993. Defining the roles of the threefold channels in iron uptake, iron oxidation and iron-core formation in ferritin: a study aided by site-directed mutagenesis. *Biochem J,* 296 (Pt 3), 721-8.

TREFFRY, A., ZHAO, Z., QUAIL, M. A., GUEST, J. R. & HARRISON, P. M. 1997. Dinuclear center of ferritin: studies of iron binding and oxidation show differences in the two iron sites. *Biochemistry,* 36, 432-41.

TSOU, C. C., CHIANG-NI, C., LIN, Y. S., CHUANG, W. J., LIN, M. T., LIU, C. C. & WU, J. J. 2008. An iron-binding protein, Dpr, decreases hydrogen peroxide stress and protects Streptococcus pyogenes against multiple stresses. *Infect Immun,* 76, 4038-45.

TSURUTA, O., YOKOYAMA, H. & FUJII, S. 2012. A new crystal lattice structure of Helicobacter pylori neutrophil-activating protein (HP-NAP). *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 68, 134-40.

UCHIDA, H., BERLINER, L. J. & KLAPPER, M. H. 1970. Three new iron-ligand complexes of human methemoglobin A. *J Biol Chem,* 245, 4606-11.

UCHIDA, M., KLEM, M. Â. T., ALLEN, M., SUCI, P., FLENNIKEN, M., GILLITZER, E., VARPNESS, Z., LIEPOLD, L. Â. O., YOUNG, M., DOUGLAS, T. 2007. Biological Containers: Protein Cages as Multifunctional Nanoplatforms. *Advanced Materials,* 19, 1025-1042.

UCHIDA, M., FLENNIKEN, M. L., ALLEN, M., WILLITS, D. A., CROWLEY, B. E., BRUMFIELD, S., WILLIS, A. F., JACKIW, L., JUTILA, M., YOUNG, M. J. & DOUGLAS, T. 2006. Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *J Am Chem Soc,* 128, 16626-33.

VALERO, E., TAMBALO, S., MARZOLA, P., ORTEGA-MUNOZ, M., LOPEZ-JARAMILLO, F. J., SANTOYO-GONZALEZ, F., DE DIOS LOPEZ, J., DELGADO, J. J., CALVINO, J. J., CUESTA, R., DOMINGUEZ-VERA, J. M. & GALVEZ, N. 2011. Magnetic nanoparticles--templated assembly of protein subunits: a new platform for carbohydrate-based MRI nanoprobes. *J Am Chem Soc,* 133, 4889-95.

VOHRADSKY, J., LI, X. M., DALE, G., FOLCHER, M., NGUYEN, L., VIOLLIER, P. H. & THOMPSON, C. J. 2000. Developmental control of stress stimulons in Streptomyces coelicolor revealed by statistical analyses of global gene expression patterns. *J Bacteriol,* 182, 4979-86.

WAHLGREN, W. Y., OMRAN, H., VON STETTEN, D., ROYANT, A., VAN DER POST, S. & KATONA, G. 2012. Structural characterization of bacterioferritin from Blastochloris viridis. *PLoS One,* 7, e46992.

WANG, Z., LI, C., ELLENBURG, M., SOISTMAN, E., RUBLE, J., WRIGHT, B., HO, J. X. & CARTER, D. C. 2006. Structure of human ferritin L chain. *Acta Crystallogr D Biol Crystallogr,* 62, 800-6.

WARDESKA, J. G., VIGLIONE, B. & CHASTEEN, N. D. 1986. Metal ion complexes of apoferritin. Evidence for initial binding in the hydrophilic channels. *J Biol Chem,* 261, 6677-83.

WIEDENHEFT, B., MOSOLF, J., WILLITS, D., YEAGER, M., DRYDEN, K. A., YOUNG, M. & DOUGLAS, T. 2005. An archaeal antioxidant: characterization of a Dps-like protein from Sulfolobus solfataricus. *Proc Natl Acad Sci U S A,* 102, 10551-6.

WINTER, G. 2010. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography,* 43, 186-190.

WOLF, S. G., FRENKIEL, D., ARAD, T., FINKEL, S. E., KOLTER, R. & MINSKY, A. 1999. DNA protection by stress-induced biocrystallization. *Nature,* 400, 83-5.

WONG, S. G., ABDULQADIR, R., LE BRUN, N. E., MOORE, G. R. & MAUK, A. G. 2012. Fe-haem bound to Escherichia coli bacterioferritin accelerates iron core formation by an electron transfer mechanism. *Biochem J,* 444, 553-60.

YAMASHITA, I., IWAHORI, K. & KUMAGAI, S. 2010. Ferritin in the field of nanodevices. *Biochim Biophys Acta,* 1800, 846-57.

YANG, X. & CHASTEEN, N. D. 1996. Molecular diffusion into horse spleen ferritin: a nitroxide radical spin probe study. *Biophysical Journal,* 71, 1587-1595.

YANG, X. & CHASTEEN, N. D. 1999. Ferroxidase activity of ferritin: effects of pH, buffer and Fe(II) and Fe(III) concentrations on Fe(II) autoxidation and ferroxidation. *Biochem J,* 338 (Pt 3), 615-8.

YAO, H., JEPKORIR, G., LOVELL, S., NAMA, P. V., WEERATUNGA, S., BATTAILE, K. P. & RIVERA, M. 2011. Two distinct ferritin-like molecules in Pseudomonas

aeruginosa: the product of the bfrA gene is a bacterial ferritin (FtnA) and not a bacterioferritin (Bfr). *Biochemistry,* 50, 5236-48.

YASMIN, S., ANDREWS, S. C., MOORE, G. R. & LE BRUN, N. E. 2011. A new role for heme, facilitating release of iron from the bacterioferritin iron biomineral. *J Biol Chem,* 286, 3473-83.

YOKOYAMA, H., TSURUTA, O., AKAO, N. & FUJII, S. 2012. Crystal structure of Helicobacter pylori neutrophil-activating protein with a di-nuclear ferroxidase center in a zinc or cadmium-bound form. *Biochem Biophys Res Commun,* 422, 745-50.

YOSHIZAWA, K., MISHIMA, Y., PARK, S. Y., HEDDLE, J. G., TAME, J. R., IWAHORI, K., KOBAYASHI, M. & YAMASHITA, I. 2007. Effect of N-terminal residues on the structural stability of recombinant horse L-chain apoferritin in an acidic environment. *J Biochem,* 142, 707-13.

ZANOTTI, G., PAPINUTTO, E., DUNDON, W., BATTISTUTTA, R., SEVESO, M., GIUDICE, G., RAPPUOLI, R. & MONTECUCCO, C. 2002. Structure of the neutrophil-activating protein from Helicobacter pylori. *J Mol Biol,* 323, 125-30.

ZETH, K. 2012. Dps biomineralizing proteins: multifunctional architects of nature. *Biochem J,* 445, 297-311.

ZETH, K., OFFERMANN, S., ESSEN, L. O. & OESTERHELT, D. 2004. Iron-oxo clusters biomineralizing on protein surfaces: structural analysis of Halobacterium salinarum DpsA in its low- and high-iron states. *Proc Natl Acad Sci U S A,* 101, 13780-5.

ZHANG, Y., FU, J., CHEE, S. Y., ANG, E. X. & ORNER, B. P. 2011. Rational disruption of the oligomerization of the mini-ferritin, E. coli DPS through protein-protein interface mutation. *Protein Sci.*

ZHANG, Y. & ORNER, B. P. 2011. Self-assembly in the ferritin nano-cage protein superfamily. *Int J Mol Sci,* 12, 5406-21.

ZHANG, Y., RAUDAH, S., TEO, H., TEO, G. W., FAN, R., SUN, X. & ORNER, B. P. 2010. Alanine-shaving mutagenesis to determine key interfacial residues governing the assembly of a nano-cage maxi-ferritin. *J Biol Chem,* 285, 12078-86.

ZHEN, Z., TANG, W., CHEN, H., LIN, X., TODD, T., WANG, G., COWGER, T., CHEN, X. & XIE, J. 2013. RGD-Modified Apoferritin Nanoparticles for Efficient Drug Delivery to Tumors. *ACS Nano,* 7, 4830-7.

ZHENG, W. J., HU, Y. H. & SUN, L. 2011. The two Dps of Edwardsiella tarda are involved in resistance against oxidative stress and host infection. *Fish Shellfish Immunol.*